Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance

Phillip A. Schwartz[a], Petr Kuzmic[b], James Solowiej[a], Simon Bergqvist[a], Ben Bolanos[c], Chau Almaden[a], Asako Nagata[c], Kevin Ryan[c], Junli Feng[c], Deepak Dalvie[d], John C. Kath[c], Meirong Xu[a], Revati Wani[a], and Brion William Murray[a,1]

[a]Oncology Research Unit, [c]Worldwide Medicinal Chemistry, and [d]Pharmacokinetics and Drug Metabolism, Pfizer Worldwide Research and Development, La Jolla, Pfizer Inc., San Diego, CA 92121; and [b]Research and Development, BioKin Ltd., Watertown, MA 02472

Edited* by Napoleone Ferrara, University of California, San Diego, La Jolla, CA, and approved November 25, 2013 (received for review July 23, 2013)

Covalent inhibition is a reemerging paradigm in kinase drug design, but the roles of inhibitor binding affinity and chemical reactivity in overall potency are not well-understood. To characterize the underlying molecular processes at a microscopic level and determine the appropriate kinetic constants, specialized experimental design and advanced numerical integration of differential equations are developed. Previously uncharacterized investigational covalent drugs reported here are shown to be extremely effective epidermal growth factor receptor (EGFR) inhibitors (k_{inact}/K_i in the range 10^5–10^7 $M^{-1}s^{-1}$), despite their low specific reactivity ($k_{inact} \leq 2.1 \times 10^{-3}$ s^{-1}), which is compensated for by high binding affinities ($K_i < 1$ nM). For inhibitors relying on reactivity to achieve potency, noncovalent enzyme–inhibitor complex partitioning between inhibitor dissociation and bond formation is central. Interestingly, reversible binding affinity of EGFR covalent inhibitors is highly correlated with antitumor cell potency. Furthermore, cellular potency for a subset of covalent inhibitors can be accounted for solely through reversible interactions. One reversible interaction is between EGFR-Cys_{797} nucleophile and the inhibitor's reactive group, which may also contribute to drug resistance. Because covalent inhibitors target a cysteine residue, the effects of its oxidation on enzyme catalysis and inhibitor pharmacology are characterized. Oxidation of the EGFR cysteine nucleophile does not alter catalysis but has widely varied effects on inhibitor potency depending on the EGFR context (e.g., oncogenic mutations), type of oxidation (sulfinylation or glutathiolation), and inhibitor architecture. These methods, parameters, and insights provide a rational framework for assessing and designing effective covalent inhibitors.

cysteine oxidation | protein kinase | signaling | capture period | warhead interactions

Receptor tyrosine kinases, such as the epidermal growth factor receptor (EGFR) tyrosine kinase, catalyze protein phosphorylation reactions to trigger signaling networks. Oncogenic activating mutations of EGFR lead to aberrant signaling for a subpopulation (10–30%) of nonsmall cell lung cancer patients (1). These mutations reside primarily in two regions of the EGFR catalytic domain [namely, the in-frame deletion mutations (e.g., Del746-750) preceding the N-terminal Cα-helix (exon 19) and the C-terminal activation loop L858R mutation (exon 21)] (2). Patients harboring these activating mutations usually respond to reversible ATP competitive drugs (e.g., erlotinib and gefitinib), but their effectiveness is limited by the emergence of drug resistance, in part, through an additional active site mutation (T790M and gatekeeper residue) in 50% of the responsive patients (3).

A second generation of drug discovery dating back to the 1990s resulted in inhibitors that incorporate a chemically reactive Michael Acceptor (MA) electrophile (warhead) to target a cysteine nucleophile (EGFR-Cys_{797}) in the hinge region of the ATP binding cleft (4). The ensuing 1,4-conjugate addition reaction of these inactivators results in an irreversible covalent adduct (Fig. 1*A*); hence, the term covalent inhibitors is used (5). To date, clinical trials of covalent EGFR inhibitors have produced mixed results (6, 7). The first covalent drug (CI-1033) did not proceed beyond early clinical studies. The next series of covalent inhibitors has advanced to phase III studies or are Food and Drug Administration-approved (3) [dacomitinib, PF-00299804 (8); neratinib, HKI-272 (9); afatinib, BIBW-2992 (10)]. Emerging clinical evidence indicates that these drugs can have superior clinical performance relative to reversible drugs but are also limited by the emergence of drug resistance (11, 12).

Covalent inhibition has reemerged as a protein kinase drug design strategy for a number of reasons (13, 14). The scope of the approach has recently expanded beyond the 11 EGFR-related kinases (e.g., HER2 and BTK) to 193 kinases that have a cysteine exposed in other regions of the active site (15). A subset of the covalent drugs is emerging as superior to reversible drugs (11, 12), which may be because of prolonged pharmacodynamic activity, lower dose, and more complete target inhibition (3). However, this mode of inhibition carries the risk of creating an immune response to epitopes from either expected or nonspecific covalent modifications (16). Therefore, the clinical benefit can outweigh the risks for well-designed inhibitors that maximize the benefit of reactivity and minimize its liability. To date, the distinct and separate contributions of noncovalent binding affinity and chemical reactivity to overall potency for the covalent EGFR investigational drugs have not been defined (17).

Significance

Covalent kinase inhibition strategies are reemerging, but critical gaps in the understanding of molecular determinants of potency still persist. A kinetic approach is developed to describe the components of overall inhibitor potency (reversible binding and chemical reactivity). Detailed kinetic descriptions of EGFR covalent drugs are provided. Reversible interactions of covalent inhibitors are found to be essential to biochemical and cellular potency. A dynamic linkage between available affinity and necessary reactivity is proposed. Cysteine oxidation is an emerging type of posttranslational modification. Specific oxidation of the EGF receptor cysteine nucleophile causes highly variable effects on inhibitor potency. Two mechanisms of drug resistance are identified (reversible cysteine–inhibitor warhead interactions and specific cysteine oxidation) as well as a rational framework for understanding and designing covalent inhibitors.

Author contributions: P.A.S., J.S., S.B., A.N., and B.W.M. designed research; P.A.S., J.S., S.B., B.B., C.A., A.N., K.R., J.F., D.D., and M.X. performed research; P.K. and J.C.K. contributed new reagents/analytic tools; P.A.S., P.K., S.B., B.B., A.N., R.W., and B.W.M. analyzed data; and B.W.M. wrote the paper.

Conflict of interest statement: All authors are employed by the companies designated in the affiliations.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: brion.murray@pfizer.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1313733111/-/DCSupplemental.

BIOCHEMISTRY

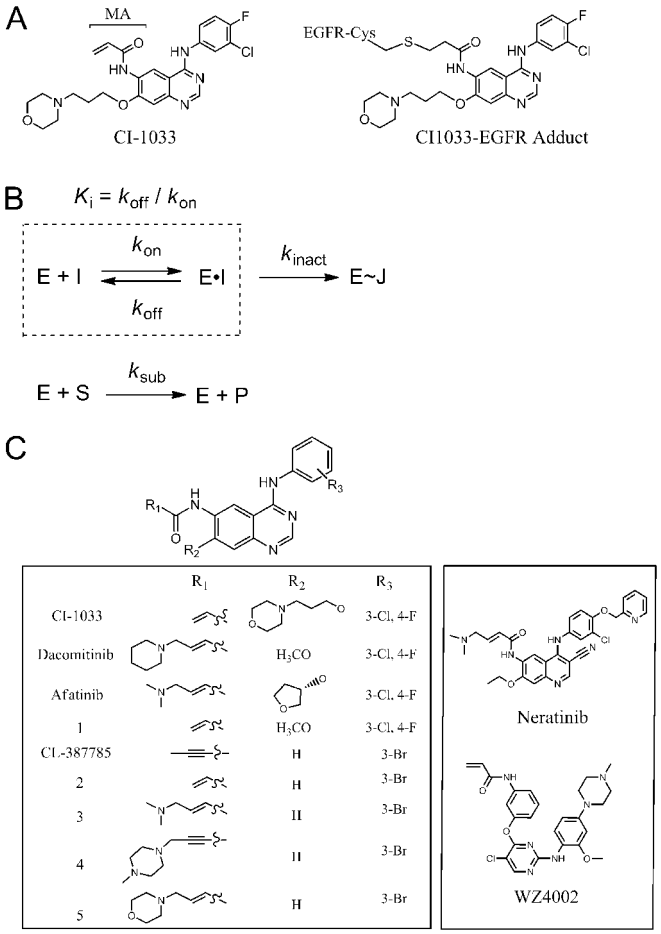


Fig. 1. (*A*) Chemical mechanisms of irreversible enzyme inhibition. Representative covalent inhibitor with reactive MA (bracket) and the resulting EGFR adduct. (*B*) The postulated kinetic mechanism for two-step covalent inhibition under the special experimental conditions where the Michaelis constant for the peptide substrate, S, is very much lower than the corresponding Michaelis constant $K_{m,Pep}$. The dashed box represents the rapid equilibrium approximation for inhibitor binding and dissociation. (*C*) Structures of EGFR covalent inhibitors investigated in this report.

In biochemical assessments of these highly potent drugs, the kinetic analysis is difficult, because deriving an exact algebraic solution is "hopelessly complex" (18). Nonetheless, defining the components of covalent inhibitor potency is important, because unlike other enzyme classes (e.g., proteases), the deep kinase active site cleft facilitates high-affinity inhibitor binding; thus, chemical reactivity can be rationally incorporated.

There is a growing appreciation that oxidation of cysteine residues affects signaling networks (19–21), including a report that EGFR is oxidized by hydrogen peroxide (H_2O_2) at Cys_{797} (the reactive nucleophile) (22, 23). Because oxidation fundamentally affects the chemical properties of the cysteine thiol by transforming it to either a highly polar oxo-acid or a bulky glutathione adduct, the EGFR active site topography and conformation may be affected. Therefore, oxidation of the nucleophilic Cys_{797} thiol has the potential to alter catalytic properties as well as covalent inhibitor potency and possibly, drug resistance. Taken together, the components of covalent inhibitor potency are critical to understanding biological effects as well as facilitating rational drug design.

Results

Defining the Covalent and Noncovalent Contributions to Overall Inhibitor Potency. The concentration of EGFR peptide substrate $[Pep]_0$ is very much lower than the corresponding Michaelis constant $K_{m,Pep}$, which makes it necessary to invoke a truncated hit-and-run (E + S → E + P) mechanistic model (Fig. 1*B* and *SI Appendix*, sections 1.1 and 1.2 show the derivation). If the peptide solubility was not a limitation, other kinetic regimes would be possible. The corresponding mathematical formalism as a system of simultaneous first-order ordinary differential equations is derived in *SI Appendix*, Section 3. The noncovalent K_i values are determined by two independent methods. The first method is based on the initial reaction rates analyzed by an algebraic fitting model (*SI Appendix*, section 2 and Table S1). This method relies on the fact that the initial enzyme–inhibitor complex is formed instantaneously on the timescale of the experiment. This rapid equilibrium assumption is applicable for all EGFR inhibitors investigated, because the empirically determined initial rates vary strongly with the inhibitor concentration, and this variation of initial rates follows the Morrison equation (24) for tight binding inhibition (*SI Appendix*, section 2, Fig. S6, and Table S1). A second, independent method of kinetic analysis relies on the global fit of complete reaction progress curves using a suitable differential equation numerical integration fitting model (*SI Appendix*, section 3). The two sets of noncovalent K_i values, determined by two independent methods, showed very good agreement ($R^2 = 0.99$) (*SI Appendix*, section 3). The inactivation rate constant k_{inact} (Fig. 1*B*) is also determined by numerical integration approach. Thus, we separate the overall inhibitory effect into two contributing components: the strength of noncovalent binding and the chemical reactivity of the initial enzyme–inhibitor complex.

Biochemical Kinetic Benchmarking of Covalent Inhibitors to WT EGFR. With this kinetic system, the first complete kinetic description of covalent drug potencies to their original therapeutic target (WT) is now possible (Fig. 2 and Table 1). Overall, the quinazoline-based covalent drugs (dacomitinib, afatinib, and CI-1033) are extremely effective ($k_{inact}/K_i = 6.3–23 \times 10^6$ $M^{-1}s^{-1}$) with high affinity ($K_i = 0.093–0.16$ nM) and low specific reactivity ($k_{inact} \leq 2.1$ ms^{-1}). As expected, these drugs are potent inhibitors of WT EGFR autophosphorylation in A549 tumor cells ($IC_{50} = 2–12$ nM). The quinolone-based investigational drug neratinib has the identical reactive substituent as afatinib, but its affinity is 50-fold weaker, with 25-fold weaker overall biochemical potency (k_{inact}/K_i). The pyrimidine-based inhibitor WZ4002 has the same reactive substituent as CI-1033 with fivefold more intrinsic chemical reactivity; however, it has 260-fold less biochemical potency. From this analysis, covalent drugs can be extremely effective

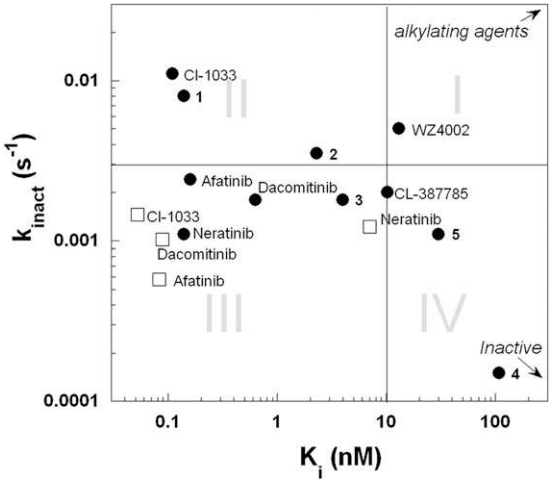


Fig. 2. Covalent drugs and inhibitors characterization based on kinetic properties with WT (□) and L858R/T790M EGFR (●): quadrant I, low affinity and high reactivity; quadrant II, high affinity and high reactivity; quandrant III, high affinity and moderate reactivity; quadrant IV, weak affinity and moderate to low reactivity.

Table 1. Kinetic analysis of covalent inhibition of EGFR-L858R/T790M WT EGFR

		EGFR-L858R/T790M				WT EGFR			
Inhibitor	GSH	K_i (nM)	k_{inact} (ms^{-1})	k_{inact}/K_i ($\mu M^{-1}s^{-1}$)	H1975 IC_{50} (nM)	K_i (nM)	k_{inact} (ms^{-1})	k_{inact}/K_i ($\mu M^{-1}s^{-1}$)	A549 IC_{50} (nM)
CI-1033	1.0	0.11 ± 0.03	11.0 ± 0.2	100 ± 20	2.3 ± 0.5	0.093 ± 0.002	2.9 ± 1.9	23 ± 1	4.9 ± 0.4
Dacomitinib	1.7	0.63 ± 0.05	1.8 ± 0.1	2.8 ± 0.3	10.3 ± 1.1	0.16 ± 0.01	1.5 ± 0.1	9.9 ± 0.8	2.5 ± 0.1
Afatinib	1.4	0.16 ± 0.03	2.4 ± 0.3	15 ± 4	7.3 ± 1.1	0.15 ± 0.01	0.9 ± 0.1	6.3 ± 0.8	11.5 ± 2.4
Neratinib	1.7	0.14 ± 0.03	1.1 ± 0.2	7 ± 2	9.4 ± 4.0	7.1 ± 0.4	1.8 ± 0.1	0.25 ± 0.01	5.2 ± 0.9
1	1.3	0.14 ± 0.07	8 ± 4	60 ± 40	6.6 ± 0.2	0.18 ± 0.01	2.3 ± 0.2	13 ± 1	5.8 ± 2.5
WZ-4002	4.8	13 ± 3	5.0 ± 0.1	0.40 ± 0.10	75 ± 25	28 ± 1	2.0 ± 0.1	0.089 ± 0.005	1,400 ± 400
CL-387785	5.0	10 ± 2	2.0 ± 0.3	0.21 ± 0.10	100 ± 7				
2	1.9	2.3 ± 0.3	3.5 ± 0.6	1.5 ± 0.3	30 ± 2				
3	2.2	4.0 ± 1.0	1.8 ± 0.1	0.40 ± 0.10	210 ± 3				
4	ND	108 ± 20	1.5 ± 0.2	0.0014 ± 0.0003	6,200 ± 3,200				
5	12.4	30 ± 3	1.1 ± 0. 1	0.04 ± 0.01	850 ± 90				

Intrinsic chemical reactivity is assessed by the reactivity to GSH relative to CI-1033 (CI-1033 half-life/inhibitor half-life). Cellular potency is quantitated by inhibition of EGFR-L858R/T790M autophosphorylation in H1975 tumor cells and EGFR WT in A549 tumor cells. ND, not determined.

EGFR inhibitors, but properties other than intrinsic chemical reactivity are critical to overall potency.

Biochemical and Cellular Characterization of Covalent Inhibition of Oncogenic EGFR. To better define the molecular determinants contributing to drug resistance as well as facilitate rational drug design of unique covalent inhibitors, the double mutant EGFR-L858R/T790M is profiled with a panel of covalent inhibitors (Table 1) encompassing three molecular scaffolds, two MAs, and the original EGFR covalent inhibitor series (3-bromo-anilino-quinazoline; CL-387785; **2**–**5**) (Fig. 1*C*) (25). Molecular modeling studies provide insight into the mode of binding for selected inhibitors (*SI Appendix*, Fig. S14). Although the quinazoline-based inhibitor **1** (PF-6274484) and the pyrimidine-based inhibitor WZ4002 share a common MA, the inhibitor scaffold places it in distinct orientations to Cys_{797}. For compound **1**, the nitrogen on the quinazoline ring interacts with the hinge region similar to EGFR drugs (gefitinib, lapatinib, and erlotinib), positioning the MA β-carbon 6 Å from Cys_{797}. WZ4002 binds to the hinge with the nitrogen on the pyrimidine ring, whereas the NH and phenoxy linkers position the MA β-carbon 3 Å from Cys_{797}. Chemical reactivity of the covalent inhibitors to the glutathione (GSH) thiol is used to assess nonenzymatic, intrinsic reactivity of these compounds (Table 1). Thus, the panel of covalent inhibitors covers a range of chemical properties and binding interactions.

Because covalent bond formation is thought to be critical, we explored the contribution of chemical reactivity to overall potency. The nonenzymatic reactivity to GSH varies only 12-fold across the inhibitor panel, whereas the specific enzymatic reactivity (k_{inact}) varies 73-fold (Table 1). There is no significant correlation between intrinsic GSH reactivity and k_{inact} ($R^2 = 0.13$), which is consistent with the EGFR architecture optimizing the reaction. There is only a moderate correlation of k_{inact} with inhibition of EGFR-L858R/T790M autophosphorylation in H1975 tumor cells ($R^2 = 0.60$) (*SI Appendix*, Fig. S10, circles). In contrast, the covalent inhibitor binding affinities (K_i) vary 750-fold across the panel (Table 1). Reversible binding affinity (K_i) correlates more strongly with cellular potency ($R^2 = 0.89$) (*SI Appendix*, Fig. S10, triangles). Overall biochemical potency as measured by the k_{inact}/K_i ratio is mostly strongly correlated with cellular potency ($R^2 = 0.95$) (*SI Appendix*, Fig. S10, squares). A broader analysis of 154 chemically diverse, cell-permeable covalent inhibitors encompassing six distinct core structures also shows a strong correlation of binding affinity with cellular potency (*SI Appendix*, Fig. S11). Inhibitor affinities are also determined for the oncogenic mutant EGFR-L858R (*SI Appendix*, Table S5). The quinazoline-based covalent drugs are found to have high binding affinities for EGFR-L858R ($K_i = 0.4$–0.7 nM), whereas the binding of pyrimidine-based WZ4002 is significantly weaker ($K_i = 13$ nM). Again, noncovalent binding affinities to EGFR-L858R are well-correlated with inhibiting EGFR-L858R autophosphorylation in tumor cells (H3255). We conclude that the initial noncovalent binding interactions leading to the formation of the initial enzyme–inhibitor complex make critically important contributions to cellular potency.

Nonreactive Analogs Reveal MA Contribution to Reversible Affinity. Nonreactive analogs of covalent inhibitors provide insight into the initial noncovalent interactions of covalent inhibitors, because the analogs have identical binding interactions, except for the fact that the MA has been fully reduced to an amide moiety (*SI Appendix*, Fig. S17). Surface plasmon resonance is used to measure the association and dissociation rate constants as well as the equilibrium binding constants of the reversible analogs of CI-1033 (**7**), dacomitinib (**8**), compound **1** (**9**), and WZ4002 (WZ4003) (*SI Appendix*, Tables S7 and S8). The dissociation equilibrium constants K_d for reversible analogs are compared with the biochemical inhibition constants K_i for the corresponding irreversible compounds. In all cases, the reversible (reduced) analogs seem to bind significantly more weakly compared with their reactive counterparts. In the case of WT EGFR, the smallest difference ($K_d/K_i = 5$) is seen for compound **7** vs. CI-1033, whereas the largest difference ($K_d/K_i = 22$) is seen for WZ4003 vs. WZ4002. The remaining two pairs of analogs in *SI Appendix*, Table S7 differ approximately by an order of magnitude ($K_d/K_i = 10$). These differences are more pronounced with EGFR-L858R/T790M, where the K_d/K_i ratio ranges from 24 (WZ4003 vs. WZ4002) to 420 (compounds **9** vs. **1**). The biochemical K_i values for the nonreactive analogs are in good agreement with their biophysical dissociation constant K_d. These results show that the MA moiety contributes significantly to the noncovalent binding affinity. The contribution of the MA to affinity is confirmed by results of cell-based assays (*SI Appendix*, Tables S5 and S6). The nonreactive analogs were evaluated to full-length, endogenous EGFR autophosphorylation in tumor cells. For example, the IC_{50} for the double mutant EGFR-L858R/T790M autophosphorylation inhibition in H1975 cells was 2.3 nM for CI-1033 but 1,800 nM (an 800-fold difference) for the corresponding nonreactive analog (compound **7**). Even more dramatically, whereas WZ4002 exhibited ~80 nM cellular IC_{50} in inhibiting both double mutant (EGFR-L858R/T790M) and single mutant (EGFR-L858R) autophosphorylation, the corresponding nonreactive analog is entirely inactive in cell-based assays. In contrast, nonreactive analogs of CI-1033 (**7**) and dacomitinib (**8**) are potent inhibitors of WT EGFR in A549 tumor cells: $IC_{50} = 16 \pm 2$ nM (**7**) and $IC_{50} = 13$ nM (**8**). The afatinib analog **6** has similar H3255 tumor cell potency (less than threefold difference compared with afatinib). Therefore, reversible MA interactions are highly variable and can contribute significantly to not only biochemical binding affinity but also, cellular potency.

Oxidation of EGFR-Cys_{797} Differentially Affects Catalysis and Inhibitor Potency. The effect of specific cysteine oxidation on inhibitor pharmacology has not been sufficiently described in the literature.

EGFR-L858R and EGFR-L858R/T790M proteins are selectively oxidized with either H_2O_2 or oxidized glutathione. Intact mass analysis reveals that a single mass shift occurs, which is consistent with either sulfinylation (-SO_2H) or S-glutathiolation. Tandem MS analysis encompassing all oxidizable residues (cysteine and methionine; 97% overall amino acid coverage) confirms that oxidation occurs exclusively at Cys_{797} (*SI Appendix*, Figs. S12 and S13). The resulting oxidized proteins are highly active, with no major change in catalytic parameters K_m and k_{cat} (*SI Appendix*, Table S4). These reagents enable us to investigate inhibitor interactions with specifically oxidized EGFR-Cys_{797}. Inhibitor binding affinities for specifically S-glutathiolated or sulfinylated EGFR-Cys_{797} are determined and compared with corresponding unoxidized forms (Fig. 3, Table 1, and *SI Appendix*, Tables S5 and S6). S-glutathiolation has the smallest effect on the reversible quinazoline drugs: 3- to 31-fold weaker affinity for EGFR-L858R (Fig. 3*A*) and 2- to 19-fold weaker affinity for EGFR-L858R/T790M (Fig. 3*B*). Quinazoline covalent inhibitors have moderately less affinity to S-glutathiolated EGFR-L858R (11- to 30-fold) (Fig. 3*C*), with much weaker affinity for S-glutathiolated EGFR-L858R/T790M (80- to 260-fold) (Fig. 3*D*). The pyrimidine inhibitor WZ4002 is a weak inhibitor of S-glutathiolated EGFR proteins (K_i > 1 μM; 170- to 400-fold less affinity) (Fig. 3*C*). EGFR-Cys_{797} sulfinylation has a distinct pharmacological profile relative to S-glutathiolation. The pyrimidine-based WZ4002 affinity is highly affected by sulfinylation of EGFR-L858R (1,100-fold) (Fig. 3*C*) and EGFR-L858R/T790M (110-fold) (Fig. 3*D*), resulting in ineffective inhibition (K_i = 1–10 μM). EGFR-L858R sulfinylation reduces reversible affinities of covalent drugs modestly (up to 12-fold) (Fig. 3*C*). Larger effects for sulfinylated EGFR-L858R/T790M (3- to 100-fold) (Fig. 3*D*) are observed when the affinity loss correlates with the complexity of the MA structure. For example, undecorated MAs (CI-1033, **1**, and WZ4002) have large affinity losses (110- to 390-fold), whereas those inhibitors with elaborate MAs (dacomitinib and afatinib) are less affected (4- to 13-fold). Taken together, the types of EGFR-Cys_{797} oxidation and inhibitor structure can profoundly alter inhibitor affinity to specifically oxidized EGFR.

Discussion

An impediment to understanding covalent drug potency is the use of overly simplified inhibitor analysis (e.g., IC_{50}) or incorrect kinetic analysis, which obscures the distinct molecular determinants (i.e., noncovalent binding affinity and chemical reactivity) that contribute separately to the overall inhibitor potency to the protein target (3, 26). In this work, we address this gap by taking a two-pronged approach. First, we optimized the experimental conditions to overcome the peptide insolubility limitation using hit-and-run experimental conditions (peptide substrate concentration very much lower than $K_{m,Pep}$ to simplify the interpretation of data and the ATP concentration very much higher than $K_{m,ATP}$ to increase assay sensitivity). Second, we introduce an advanced method of irreversible inactivation data analysis based on the numerical integration of the full system of simultaneous first-order ordinary differential equations (27). This approach allows us to efficiently define overall inhibitor potency in terms of contributions from reversible and irreversible components.

Having defined the individual determinants of inhibitory potency, we are able to organize covalent inhibitor potency space into four quadrants based on initial binding affinity and chemical reactivity (Fig. 2). Quadrant 1 contains alkylating agents that rely on intrinsic reactivity to achieve potency. Quadrant 4 is reserved for low-potency inhibitors. The challenge has been to discriminate between effective inhibitors that achieve potency by high affinity and reactivity (quadrant 2) or high affinity and moderate reactivity (quadrant 3). Covalent EGFR drugs have extremely high affinity and low reactivity to WT EGFR (quadrant 3). However, for EGFR-L858R/T790M, afatinib and CI-1033 map to quadrant 2, which indicates that the drug-resistant form of EGFR presents Cys_{797} differently.

The individual kinetic parameters can be used to characterize the molecular interactions underlying overall inhibitory potency. The K_i value is relatively simple to interpret, because it measures reversible binding affinity. In contrast, many different factors influence the k_{inact} value, including intrinsic chemical reactivity, reactant alignment, enforced local concentration, and the cysteine pK_a. To enhance the use of the microscopic rate constant k_{inact}, we introduce its inverse value ($1/k_{inact}$) and assign to it the term capture period (t_C) measured in the units of time. The typical values of k_{inact} for EGFR inhibitors in this study ranged from 1 to 10 ms^{-1}, which corresponds to t_C = 1,000 s (~15 min) and t_C = 100 s (~1.5 min), respectively. The capture period concept is analogous to the notion of reversible inhibitor residence time ($1/k_{off}$) (17), which has found use in drug design. These two values can be used in concert to facilitate rational covalent inhibitor design.

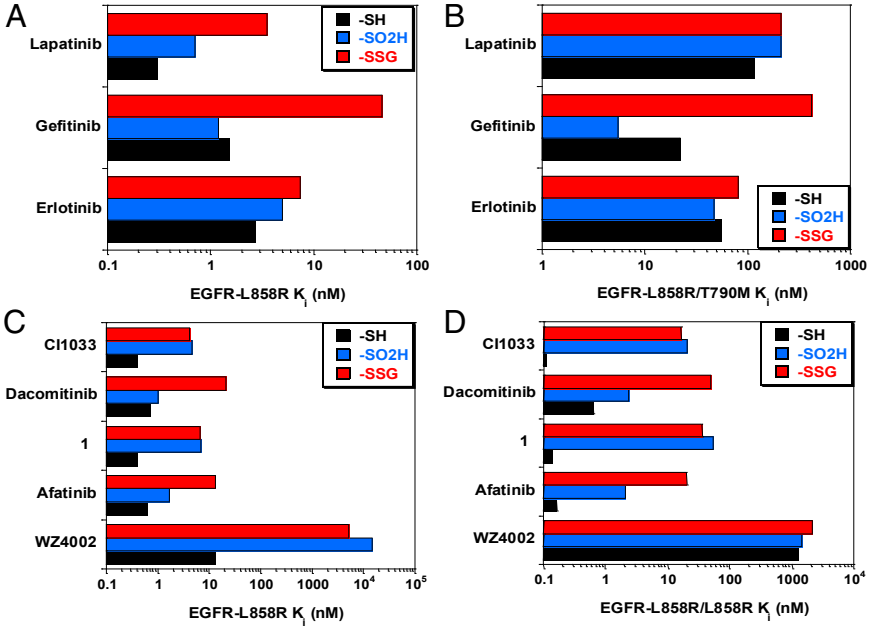


Fig. 3. Specific EGFR-Cys_{797} oxidation has differential effects on inhibitor and drug potencies dependent on the type of oxidation and EGFR mutation. Reversible drug affinity determined to different Cys_{797} oxidation states (-SH, unoxidized; -SO_2H, sulfinylated; -SSG, glutathiolated) in (*A*) L858R and (*B*) L858R/T790M. Covalent inhibitor affinity was measured to (*C*) L858R and (*D*) L858R/T790M.

Reversible EGFR Complexes with Covalent Inhibitors Are Integral to Overall Potency. The contribution of a reversible enzyme–inhibitor complex to covalent inhibitor potency is currently underappreciated. For example, covalent EGFR inhibitors have been incorrectly reported to have an infinite active site residence time (28), implying that the inhibitor never actually dissociates from the initially formed noncovalent complex. However, our analysis shows that the reversible inhibitor–EGFR complex preceding the covalent adduct formation is critical to cellular potency. Reversible inhibitor binding to EGFR affects the chemical reaction by multiple mechanisms. First, it defines the MA moiety orientation to the cysteine nucleophile. Second, it increases the effective reactant concentrations. Third, it defines the number of binding/release cycles necessary for a chemical reaction to occur. Nonproductive binding events (no reaction) should be expected for many reasons, such as misaligned reactants (protein conformations are dynamic) or cysteine nucleophile in the less-reactive thiol form. An important aspect of the reversible inhibitor complex typically overlooked is that it alone is sufficient to block enzymatic function.

MA Contributes a Key Reversible Interaction to Binding Affinity. Covalent inhibitor complexes that do not result in adduct formation may still have reversible Cys_{797}•MA interactions that can contribute to potency. Binding affinity of all covalent inhibitors can now be directly measured by enzyme kinetics to quantify the sum of all binding interactions. Nonreactive analogs have identical structures, except that they do not have the rigid, planar, π-electron–containing MA substituent. Therefore, the difference in covalent inhibitor and reversible analog affinities could result from reversible Cys_{797}•MA interactions. Although the difference in affinities could be caused by steric interactions of the flexible, nonreactive substituent, it is probably a minor component, because the nonreactive quinazoline inhibitors have EGFR-L858R affinities similar to covalent inhibitors. A precedent for S–π or SH–π interactions is found in both model systems (e.g., H_2S–benzene) (29) and protein structures (30) with the capability of contributing significant binding energy (e.g., 2.6 kcal/mol). We find that the contribution of the reversible Cys_{797}•MA interactions to potency can be large and depend on both inhibitor properties and EGFR context. Therefore, covalent inhibitor affinity can be derived, in part, by reversible MA interactions with EGFR-Cys_{797}.

Expression of Covalent Inhibitor Potency in Tumor Cells. We have shown that overall covalent inhibitor biochemical potency is derived from contributions of reversible binding affinity and covalent adduct formation. Cellular analysis of covalent inhibitors is complicated, because as they achieve higher affinity, the need for covalent adduct formation becomes less important. Contributing to the complexity is that in cells, protein kinases are subject to temporal regulation by multiple cellular mechanisms (e.g., activation, trafficking, recycling, and degradation) (31, 32). The translation of biochemical interactions to cellular potency can now be more finely dissected with the array of tools available. Quinazoline covalent inhibitors have extremely high reversible affinity. Nonreactive analogs of CI-1033 and dacomitinib are still very effective inhibitors of WT ($K_d \sim 1$ nM, A549 $IC_{50} \sim 10$ nM). Although the covalent adduct formation does occur, the expression of cellular potency can be accounted for solely by reversible interactions, which include those interactions with the MA. The cellular potency gained through the inclusion of a reactive MA correlates with the enhanced reversible affinity of the covalent inhibitor. In the context of L858R/T790M mutations, there is significantly less affinity for the reversible analogs of dacomitinib and CI-1033 (K_d = 20–50 nM), and thus, the MA contributes to reversible affinity by 30- to 400-fold. In contrast, other inhibitors may rely more heavily on covalent adduct formation to achieve potent inhibition. The pyrimidine inhibitor WZ4002 has a different Cys_{797}•MA orientation (*SI Appendix*, Fig. S14). and without a reactive MA, there is no detectable cellular potency. Collectively, understanding the contribution of reversible binding interactions is critical to defining the molecular interactions resulting in cellular potency.

Potential Mechanisms of Covalent Inhibitor Drug Resistance. The more fine-grained characterization of covalent inhibitor properties made possible by our approach (examining k_{inact} and K_i separately) may explain conflicting reports on covalent drug resistance. One study reports that CI-1033 elicits a C797S resistance mutation in a cellular model system and causes an ∼200-fold reduction in cellular potency (33). Another study reports that a related covalent inhibitor CL-387785 does not elicit a C797S resistance mutation in H1975 cells and causes only a fourfold loss of cellular potency (34). We find that disrupting the reversible Cys_{797}•MA interaction of CI-1033 by eliminating the acrylamide MA results in 250-fold weaker affinity for EGFR-L858R/T790M and an 800-fold reduction of cellular potency. This observation indicates a reliance of the reversible Cys_{797}•MA interaction, and CI-1033 drug resistance by C797S mutation seems likely. Replacement of the CL-387785 MA propyl alkynamide by a nonreactive methyl group results in only a 10-fold reduction in EGFR-L858R/T790M affinity ($K_i = 110 \pm 2$ nM), which predicts a lower reliance on a Cys_{797}•MA interaction and a lower likelihood of an additional C797S drug resistance mutation. Our analysis reinforces the concept that specific interactions of covalent inhibitors with EGFR-Cys_{797} are critical to the development of drug resistance.

Implications of Specific Oxidation of the Reactive Cysteine Nucleophile. Recently, the WT EGFR cysteine nucleophile EGFR-Cys_{797} was shown to be oxidized by H_2O_2 (22). However, the ultimate oxidation state was not elucidated, because a transient intermediate oxidation state was trapped (sulfenic acid and -SOH). We now show that the EGFR-Cys_{797} thiol can be stably oxidized to either the sulfinic acid ($-SO_2H$) or the S-glutathiolated adduct while retaining catalytic activity. Because EGFR-Cys_{797} is not expected to be an effective nucleophile when oxidized, the impact on covalent inhibition seems straightforward, because covalent bond formation is now precluded, which results in significant loss of overall effectiveness. However, covalent inhibitors can still undergo reversible association with the altered active site topography of oxidized EGFR. S-glutathiolation introduces substantial steric bulk, which can have both steric and conformational effects on inhibitor binding. H_2O_2 oxidation of the Cys_{797} thiol to sulfinic acid (Cys_{797}-SO_2H) introduces little additional steric bulk but substantially changes the residue's polarity. We identified certain effects of different oxidation states that are specific to a given inhibitor and the given EGFR mutant. Reversible EGFR-directed drugs, such as gefitinib, are not negatively affected by sulfinylation, whereas pyrimidine-based covalent inhibitors have low binding affinity under identical circumstances. Covalent quinazoline inhibitors display mixed effects, which may be caused by favorable sulfinylated cysteine residue interactions for those inhibitors with a well-aligned basic group substituent (dacomitinib, afatinib, and gefitinib) (*SI Appendix*, Figs. S15 and S16). From this biochemical analysis, the effect of cysteine oxidation can be highly variable depending on both the type of oxidation and specific inhibitor active site interactions.

Conclusions

A general approach is developed to characterize the component biochemical processes of covalent inhibitor potency. With this approach, we are able to determine both the equilibrium dissociation constant of the initial noncovalent complex and the chemical reactivity of the warhead moiety. Factoring the overall inhibitory effect into reversible and irreversible components enables a deeper understanding and description of this reemerging inhibition modality. Reversible interactions of covalent inhibitors with EGFR, including Cys_{797}•MA, are shown to be essential to both biochemical and cell potencies. In addition, specific cysteine oxidation has been identified here as a possible drug resistance mechanism. Taken together, we present a rational framework for understanding and optimizing covalent enzyme inhibitors, which encompasses the inhibitor structure, binding interactions, and the cysteine nucleophile.

Materials and Methods

Additional details are in *SI Appendix*.

Synthesis of Sulfinylated EGFR-Cys$_{797}$. EGFR-L858R (40 μM) or EGFR-L858R/T790M (40 μM) was incubated (15 min at 23 °C) in a 120-μL reaction [100 μM H_2O_2, 60 mM NaCl, 15 mM KCl, 0.01% Tween-20, 10% (vol/vol) glycerol, 100 mM $(NH_4)HCO_3$, pH 9.2] and loaded onto 3× Zeba spin desalting columns equilibrated with 150 mM NaCl, 10% glycerol, 0.01% Tween-20, and 2 mM DTT in 50 mM Hepes (pH 7.5) at 4 °C. Sulfinylation at Cys$_{797}$ was >90% by MS.

Synthesis of Glutathiolated EGFR-Cys$_{797}$. EGFR-L858R (8 μM) or EGFR-L858R/T790M (8 μM) was incubated for 185 min at 23 °C in a 600-μL reaction (2 mM oxidized glutathione, 100 mM NaCl, 0.01% Tween-20, 10% glycerol, 25 mM Hepes, pH 8.5). Reaction aliquots (0.20 mL) were concentrated fivefold (3× Microcon spin concentrators, 14,000 × *g* for 15 min at 23 °C) to 40 μL and loaded onto 3× Zeba spin desalting columns (100 mM NaCl, 10% glycerol, 0.01% Tween-20, 25 mM Hepes, pH 7.5) at 4 °C.

EGFR Kinase Activity Assays. A coupled enzymatic spectrometric assay measuring ADP production is used to determine catalytic constants and reversible inhibitor potency (35). Covalent inhibitor analysis uses an Omnia continuous fluorometric assay with a Y12 tyrosine phosphoacceptor peptide [Ac-EEEEYI(cSx)IV-NH_2; Invitrogen] (36).

Intrinsic Chemical Reactivity Assay. Inhibitor reactivity with GSH is assessed by monitoring inhibitor loss during the reaction: 0.1 M phosphate buffer (pH 7.4), 0.1 μM compounds, and 5 mM GSH at 37 °C. The intrinsic reactivity is reported as the ratio of half-lives: inhibitor to CI-1033 (CI-1033 $t_{1/2} = 10$ min).

EGFR Cellular Autophosphorylation ELISA. For tumor cell lines, inhibitors were incubated with cells (H3255, L858R; NCI-H1975, L858R/T790M; A549, WT) for 2 h. PathScan Phospho-EGF Receptor (Tyr1068) Sandwich ELISA (Cell Signaling Technology) was quantitated per the manufacturer's protocol.

Mass Spectrometric Analysis of EGFR. Intact mass analysis used electrospray ionization on an Agilent 6210 time-of-flight mass spectrometer coupled to an Agilent 1200 LC. To identify oxidized residues, a Proxeon nanoLC coupled to an LTQ mass spectrometer was used on samples isolated by nondenaturing PAGE, pepsin-proteolyzed, and purified by Reprosil ProteCol Trap C18-AQ and Halo ES-C18 columnchromatography. MS-MS data are processed in Agilent Spectrum Mill rev. 4.0.

Inhibitor Modeling and Docking Methodology. Simulations used Glide in Standard Precision mode. EGFR cocrystal structures [gefitinib, Protein Data Bank (PDB) ID code 2ITZ; lapatinib, PDB ID code 1XKK; erlotinib, PDB ID code 1M17] optimized and minimized for the docking simulation.

Analysis of Enzyme Kinetic Data. Initial reaction rates were determined by a least squares fit of the initial portion ($t_{max} < 7$ min) of progress curves to the single exponential equation. Dissociation constants of the initial noncovalent enzyme/inhibitor complex were determined by two independent methods: (*i*) from the initial reaction rates and (*ii*) from the global fit of the complete reaction progress curves. The underlying system of first-order ordinary differential equations was integrated using the LSODE algorithm (37–39).

ACKNOWLEDGMENTS. We thank Drs. Karen Maegley and Stephan Grant for helpful feedback. We also thank the Pfizer Postdoctoral Program for supporting the fellowships of P.A.S. and R.W.

1. Gazdar AF (2009) Activating and resistance mutations of EGFR in non-small-cell lung cancer: Role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene* 28(Suppl 1):S24–S31.
2. Sharma SV, Bell DW, Settleman J, Haber DA (2007) Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 7(3):169–181.
3. Schwartz PA, Murray BW (2011) Protein kinase biochemistry and drug discovery. *Bioorg Chem* 39(5–6):192–210.
4. Wissner A, Mansour TS (2008) The development of HKI-272 and related compounds for the treatment of cancer. *Arch Pharm (Weinheim)* 341(8):465–477.
5. Silverman RB (1995) Mechanism-based enzyme inactivators. *Methods Enzymol* 249: 240–283.
6. Jänne PA, et al. (2007) Multicenter, randomized, phase II trial of CI-1033, an irreversible pan-ERBB inhibitor, for previously treated advanced non small-cell lung cancer. *J Clin Oncol* 25(25):3936–3944.
7. Rixe O, et al. (2009) A randomized, phase II, dose-finding study of the pan-ErbB receptor tyrosine-kinase inhibitor CI-1033 in patients with pretreated metastatic breast cancer. *Cancer Chemother Pharmacol* 64(6):1139–1148.
8. Jänne PA, et al. (2011) Phase I dose-escalation study of the pan-HER inhibitor, PF299804, in patients with advanced malignant solid tumors. *Clin Cancer Res* 17(5):1131–1139.
9. Burstein HJ, et al. (2010) Neratinib, an irreversible ErbB receptor tyrosine kinase inhibitor, in patients with advanced ErbB2-positive breast cancer. *J Clin Oncol* 28(8):1301–1307.
10. Yang JC, et al. (2012) Afatinib for patients with lung adenocarcinoma and epidermal growth factor receptor mutations (LUX-Lung 2): A phase 2 trial. *Lancet Oncol* 13(5):539–548.
11. Kwak E (2011) The role of irreversible HER family inhibition in the treatment of patients with non-small cell lung cancer. *Oncologist* 16(11):1498–1507.
12. Singh J, et al. (2012) Superiority of a novel EGFR targeted covalent inhibitor over its reversible counterpart in overcoming drug resistance. *Med Chem Commun* 3:780–783.
13. Singh J, Petter RC, Baillie TA, Whitty A (2011) The resurgence of covalent drugs. *Nat Rev Drug Discov* 10(4):307–317.
14. Garuti L, Roberti M, Bottegoni G (2011) Irreversible protein kinase inhibitors. *Curr Med Chem* 18(20):2981–2994.
15. Leproult E, Barluenga S, Moras D, Wurtz JM, Winssinger N (2011) Cysteine mapping in conformationally distinct kinase nucleotide binding sites: Application to the design of selective covalent inhibitors. *J Med Chem* 54(5):1347–1355.
16. Nakayama S, et al. (2009) A zone classification system for risk assessment of idiosyncratic drug toxicity using daily dose and covalent binding. *Drug Metab Dispos* 37(9):1970–1977.
17. Copeland RA, Pompliano DL, Meek TD (2006) Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* 5(9):730–739.
18. Williams JW, Morrison JF, Duggleby RG (1979) Methotrexate, a high-affinity pseudosubstrate of dihydrofolate reductase. *Biochemistry* 18(12):2567–2573.
19. Paulsen CE, Carroll KS (2010) Orchestrating redox signaling networks through regulatory cysteine switches. *ACS Chem Biol* 5(1):47–62.
20. Poole LB, Nelson KJ (2008) Discovering mechanisms of signaling-mediated cysteine oxidation. *Curr Opin Chem Biol* 12(1):18–24.
21. Filosto S, et al. (2011) EGF receptor exposed to oxidative stress acquires abnormal phosphorylation and aberrant activated conformation that impairs canonical dimerization. *PLoS One* 6(8):e23240.
22. Paulsen CE, et al. (2012) Peroxide-dependent sulfenylation of the EGFR catalytic site enhances kinase activity. *Nat Chem Biol* 8(1):57–64.
23. Truong TH, Carroll KS (2012) Redox regulation of EGFR signaling through cysteine oxidation. *Biochemistry* 51:9954–9965.
24. Morrison JF (1969) Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochim Biophys Acta* 185(2):269–286.
25. Tsou HR, et al. (2001) 6-Substituted-4-(3-bromophenylamino)quinazolines as putative irreversible inhibitors of the epidermal growth factor receptor (EGFR) and human epidermal growth factor receptor (HER-2) tyrosine kinases with enhanced antitumor activity. *J Med Chem* 44(17):2719–2734.
26. Copeland RA (2005) *Evaluation of Enzyme Inhibitors in Drug Discovery* (Wiley, New York).
27. Szedlacsek SE, Duggleby RG (1995) Kinetics of slow and tight-binding inhibitors. *Methods Enzymol* 249:144–180.
28. Heuckmann JM, Rauh D, Thomas RK (2012) Epidermal growth factor receptor (EGFR) signaling and covalent EGFR inhibition in lung cancer. *J Clin Oncol* 30(27):3417–3420.
29. Ringer AL, Senenko A, Sherrill CD (2007) Models of S/pi interactions in protein structures: Comparison of the H2S benzene complex with PDB data. *Protein Sci* 16(10): 2216–2223.
30. Daeffler KN, Lester HA, Dougherty DA (2012) Functionally important aromatic-aromatic and sulfur-π interactions in the D2 dopamine receptor. *J Am Chem Soc* 134(36): 14890–14896.
31. Carpenter G, Liao HJ (2009) Trafficking of receptor tyrosine kinases to the nucleus. *Exp Cell Res* 315(9):1556–1566.
32. Murphy JE, Padilla BE, Hasdemir B, Cottrell GS, Bunnett NW (2009) Endosomes: A legitimate platform for the signaling train. *Proc Natl Acad Sci USA* 106(42): 17615–17622.
33. Avizienyte E, Ward RA, Garner AP (2008) Comparison of the EGFR resistance mutation profiles generated by EGFR-targeted tyrosine kinase inhibitors and the impact of drug combinations. *Biochem J* 415(2):197–206.
34. Yu Z, et al. (2007) Resistance to an irreversible epidermal growth factor receptor (EGFR) inhibitor in EGFR-mutant lung cancer reveals novel treatment strategies. *Cancer Res* 67(21):10417–10427.
35. Murray BW, Padrique ES, Pinko C, McTigue MA (2001) Mechanistic effects of autophosphorylation on receptor tyrosine kinase catalysis: Enzymatic characterization of Tie2 and phospho-Tie2. *Biochemistry* 40(34):10243–10253.
36. Luković E, González-Vera JA, Imperiali B (2008) Recognition-domain focused chemosensors: Versatile and efficient reporters of protein kinase activity. *J Am Chem Soc* 130(38):12821–12827.
37. Hindmarsh AC (1983) ODEPACK: A systematized collection of ODE solvers. *Scientific Computing*, ed Vichnevetsky R (North-Holland, Amsterdam), pp 55–64.
38. Kuzmic P (1996) Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase. *Anal Biochem* 237(2):260–273.
39. Kuzmic P (2009) DynaFit—a software package for enzymology. *Methods Enzymol* 467:247–280.

SUPPORTING INFORMATION

Covalent EGFR Inhibitor Analysis Reveals Importance of Reversible Interactions to Potency and Mechanisms of Drug Resistance

Phillip A. Schwartz[a], Petr Kuzmič[b], James Solowiej[a], Simon Bergqvist[a], Ben Bolanos[c], Chau Almaden[a], Asako Nagata[c], Kevin Ryan[c], Junli Feng[c], Deepak Dalvie[d], John Kath[c], Meirong Xu[a], Revati Wani[a], Brion W. Murray[*,a]

[a]*Oncology Research Unit, Pfizer Worldwide Research and Development, La Jolla, California*
[b]*BioKin Ltd., Watertown, Massachusetts*
[c]*Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, La Jolla, California*
[d]*Pharmacokinetics and Drug Metabolism, Pfizer Worldwide Research and Development, La Jolla, California*

Contents

*Corresponding author

1. Substrate kinetics

This section describes the determination of kinetic properties of both the fluorescent peptide substrate "Y12", and ATP as the co-substrate, in the Omnia assay. We first demonstrate by theoretical analysis that the specificity number $k_{sub} = k_{cat}/K_m$ is well defined by the experimental data collected at low peptide substrate concentration ($[S]_0 << K_m$) provided that the enzyme's active-site concentration is known independently. We also derive a reduced "hit-and-run" mechanistic model ($E+S \rightarrow E+P$) for substrate catalysis, which can be used to auto-generate a suitable system of first-order ordinary differential equations (ODEs) under these specialized experimental conditions. Finally, we present experimental data to show that the Michaelis constant for the peptide substrate, $K_{m,Pep}$, cannot be determined because it is immeasurably high. However, we report a lower-limit estimate for $K_{m,Pep}$ and demonstrate that it is indeed significantly higher than the maximum experimentally attainable peptide substrate concentration. These experimental findings justify our truncated theoretical model for the substrate catalysis branch of the overall inhibition mechanism.

1.1. Kinetic theory under the $[S]_0 << K_m$ conditions

Algebraic analysis. Under the experimental conditions where the substrate concentration remains very much lower than the Michaelis constant, $[S]_0 << K_m$, the classic Michaelis-Menten rate Eqn (1) reduces to Eqn (2). This is because if $[S]_0 << K_m$ then $[S]_0$ can be neglected in the summation term and therefore we obtain the approximate equality $[S]_0 + K_m \approx K_m$.

$$v = [E]_0 \, k_{cat} \, \frac{[S]_0}{[S]_0 + K_m} \quad (1)$$

$$v \approx [E]_0 \, k_{cat} \, \frac{[S]_0}{K_m} \quad (2)$$

The reaction rate is defined as the rate of change in substrate concentration, taken with negative sign, because the substrate is being consumed in the enzyme-catalyzed reaction, $v \equiv= -\mathrm{d}[S]/\mathrm{d}t$. The resulting differential equation (3) can be integrated analytically after separation of variables, to obtain the integral equation (4) as the theoretical model for reaction progress.

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = -[E]_0\, k_{\mathrm{sub}}\, [S] \quad (3)$$

$$[S] = [S]_0\, \exp(-[E]_0\, k_{\mathrm{sub}}\, t) \quad (4)$$

$$k_{\mathrm{sub}} \equiv \frac{k_{\mathrm{cat}}}{K_{\mathrm{m}}} \quad (5)$$

Thus, at substrate concentrations significantly lower than the Michaelis constant, the reaction progress curve conforms to the simple exponential model, and the apparent first-order rate constant is equal to $[E]_0\, k_{\mathrm{cat}}/K_{\mathrm{m}}$. The enzyme concentration $[E]_0$ can be determined experimentally by active-site titration. Therefore, it is possible to determine the specificity number $k_{\mathrm{sub}} = k_{\mathrm{cat}}/K_{\mathrm{m}}$ simply by fitting the reaction progress to the first-order exponential equation and dividing the apparent first-order rate constant by $[E]_0$.

In conclusion, enzyme kinetic data observed under first-order conditions ($[S]_0 << K_{\mathrm{m}}$) do contain sufficient information (in the information-theoretic sense [1]) about k_{sub} as a potential fitting parameter.

Differential Equation Modeling. A number of software packages [2–4] allow the biochemist to specify the kinetic mechanism of an enzyme reaction using a symbolic notation, such as $E + S \rightleftharpoons ES \rightarrow E + P$. The software then internally derives the corresponding mathematical model for the reaction progress, as a system of first-order ordinary differential equations (ODEs). The question arises what particular symbolic notation should be used to define the time-course of an enzyme reaction following the Michaelis-Menten kinetic mechanism under the first-order experimental conditions ($[S]_0 << K_{\mathrm{m}}$).

We do require that the reaction rate is strictly proportional to the enzyme concentration. We also require that the integrated kinetic equation describes a first-order exponential. This leads to the differential equation (3), which corresponds to the symbolic term $E + S \rightarrow$ .... The remaining question is what should appear on the right-hand side. The answer lies in the requirement that the enzyme concentration should be treated as a constant in Michaelis-Menten kinetics. Therefore the free enzyme must appear on the right hand side, because only then the rate terms for the consumption and generation of E cancel out, $E + S \rightarrow E + P$, resulting in zero overall change in enzyme concentration over time. Thus we arrive at the "hit-and-run" symbolic model for enzyme catalysis, shown in *Scheme S1*.

$$\mathrm{E + S} \xrightarrow{k_{\mathrm{sub}}} \mathrm{E + P}$$

Scheme S1

In DynaFit notation [2], the "hit-and-run" model for substrate catalysis is represented as "`E + S --> E + P : ksub`". *Scheme S1* is conceptually similar to the Theorell-Chance mechanism for bisubstrate enzymes [5, p. 594] ($E + A \rightarrow EA; EA + B \rightarrow EQ + P; EQ \rightarrow E + Q$). No ternary molecular complex is postulated, even though at least one such complex is *implied*. In the "hit-and-run" mechanism, we also know that the Michaelis complex ES must be physically present. We are choosing to ignore it, so that we can produce a *practically useful* minimal model of the reaction progress.

1.2. Substrate kinetics of the "Y12" peptide

The"Y12" peptide (Ac-EEEEYI(cSx)IV-NH_2, Omnia Y Peptide 12, Invitrogen / Life Technologies) was assayed at ATP concentrations subsequently used in all inhibition assays ([ATP] = 800 μM) and at the enzyme concentration set to $[E]_0$ = 50 nM. The peptide concentration $[S]_0$ was varied from 4 μM to 12 μM stepping by 2 μM, in triplicate. Fig. S1 shows a representative set of the reaction progress curves. Reaction progress curves are nonlinear throughout the entire assay and conform to the exponential fitting model, suggesting that the Michaelis constant is very much larger than the maximum substrate concentration.

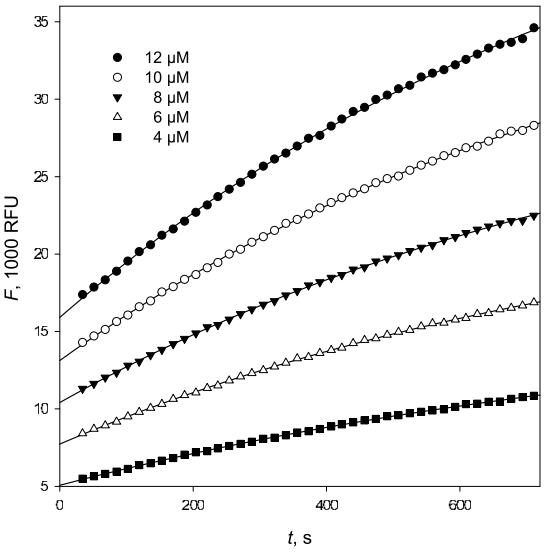


Fig. S 1: Substrate kinetics of the "Y12" peptide vs. EGFR-L858R/T790M double mutant. *Legend:* peptide concentrations used in each kinetic experiment.

All reaction progress curves were nonlinear throughout the entire time course of the assay. The progress curves conformed ideally to the exponential fitting equation (6), where F is the fluorescence intensity recorded at the reaction time t; F_0 is the baseline fluorescence (an instrument offset); A is the exponential amplitude; and k is the empirical first order rate constant. Initial reaction rates were computed from the best-fit values of

A and k by using Eqn (7). The initial rates were subsequently fit to the Michaelis-Menten rate equation (8), in an attempt to determine the Michaelis constant K_{m}.

$$F = F_0 + A\ [1 - \exp(-k\ t)] \tag{6}$$

$$v = A\ k \tag{7}$$

$$v = V_{\max} \frac{[S]_0}{[S]_0 + K_{\mathrm{m}}} \tag{8}$$

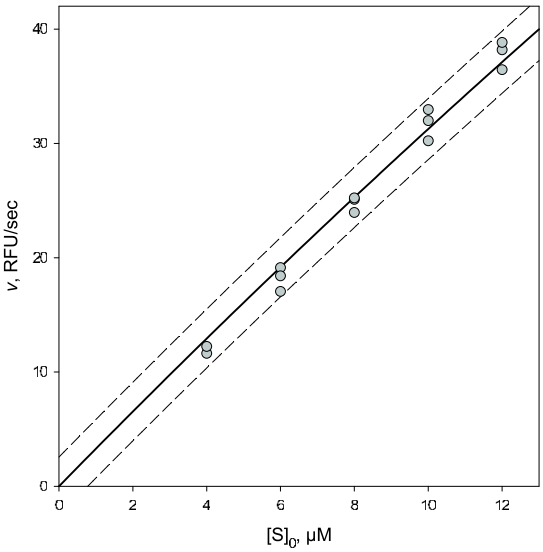


Fig. S 2: Substrate kinetics of the "Y12" peptide vs. EGFR-L858R/T790M double mutant: Circles - initial reaction rates computed from the best-fit values of exponential amplitude A and exponential rate constants k, using Eqn (7). *Solid curve* - least square model curve corresponding to Eqn (8) while keeping the Michaelis constant at the lower limit of its 95% confidence interval, K_{m} = 174 μM. *Dashed curves* - inference (prediction) bands at the 95% confidence level.

Fig. S2 shows the results of the least-squares fit of the initial reaction rates, computed from Eqn (7), to the Michaelis-Menten rate Eqn (8). The best least-squares fit values and the corresponding standard errors of model parameters $V_{\max}$ and K_{m} could not be determined. In fact, the least-squares model curve was essentially a perfect straight line through the origin, corresponding to infinite values of both $V_{\max}$ and K_{m}.

However, we were able to estimate at least to lower limit of the confidence interval for K_{m} utilizing the profile-t (likelihood profile) method of Bates & Watts [6, 7]. Using this method we determined, at the 95% confidence level, that the apparent K_{m} value of the peptide substrate, at [ATP] = 800 μM, must be larger than 174 μM. This value is very much larger than the peptide substrate concentration used in all our Omnia inhibition assays ([Pep] = 13 μM).

Thus, we conclude that in all enzyme assays we employed in this study to characterize the irreversible inhibitors of EGFR, the peptide substrate concentration [S] = 13 μM is (effectively) negligibly small when compared to the corresponding Michaelis constant. This justifies the use of the truncated ("hit-and-run") kinetic model described in *Scheme S1*, which gives rise to first-order exponential kinetics in the absence of inhibitors, as defined by Eqn (9). In Eqn (9), $[S]$ is the substrate concentration at the arbitrary reaction time t; $[S]_0$ is the initial substrate concentration at the $t = 0$; $[E]_0$ is the enzyme concentration; and k_{sub} is the second-order rate constant by definition equal to $k_{\mathrm{cat}}/K_{\mathrm{m}}$.

$$[S] = [S]_0\ \exp(-k_{\mathrm{sub}}\ [E]_0\ t) \tag{9}$$

It should be noted that under the special experimental conditions where the substrate concentration is effectively infinitely lower than the corresponding Michaelis constant ($[S]_0 << K_{\mathrm{m}}$), essentially all enzyme exists in the unbound (free) state, and thus the biochemical system behaves as of the enzyme-substrate complex were not formed at all.

1.3. Substrate kinetics of ATP

ATP was assayed at the "Y12" peptide (Ac-EEEEYI(cSx)IV-NH2, Omnia Y Peptide 12, Invitrogen / Life Technologies) concentrations subsequently used in all inhibition assays ([Pep] = 13 μM) and at the enzyme concentration set to $[\mathrm{E}]_0$ = 30 nM. ATP concentration $[\mathrm{ATP}]_0$ was varied according to the dilution series shown in the inset to Fig. S3. All assays were performed in triplicate, on separate micro-titer plates. Fig. S3 shows one of three replicated sets of the reaction progress curves.

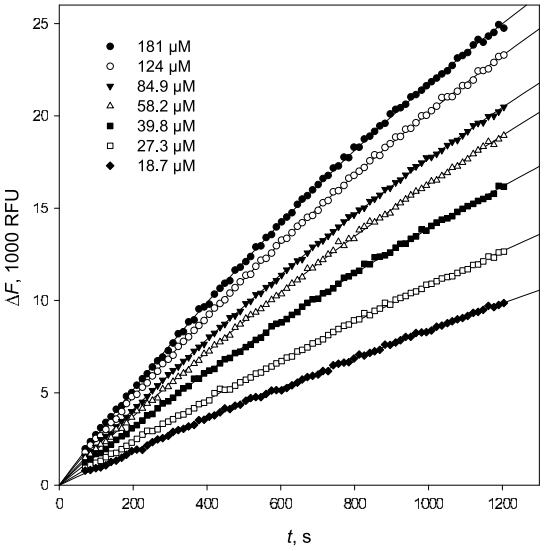


Fig. S 3: Substrate kinetics of ATP (Replicate 1 of 3) vs. EGFR-L858R/T790M double mutant: Reaction progress curves are nonlinear throughout the entire assay and conform to the exponential fitting model. *Legend:* ATP concentrations used in each kinetic experiment.

As was the case for peptide assays described above, all reaction progress curves were nonlinear throughout the entire time course of the assay. The progress curves conformed ideally to the exponential fitting equation (6). Initial reaction rates were computed from the best-fit values of A and k by using Eqn (7). The initial rates were subsequently fit to the Michaelis-Menten rate equation (8), in order to determine the Michaelis constant K_{m}. Fig. S4 shows the results of the least-squares fit of the

initial reaction rates. The best least-squares fit values and the corresponding standard errors of model parameters were $V_{\max}$ = (37.7 ± 1.5) RFU/sec and $K_{\mathrm{m,ATP}}$ = (49.1 ± 5.2) μM.

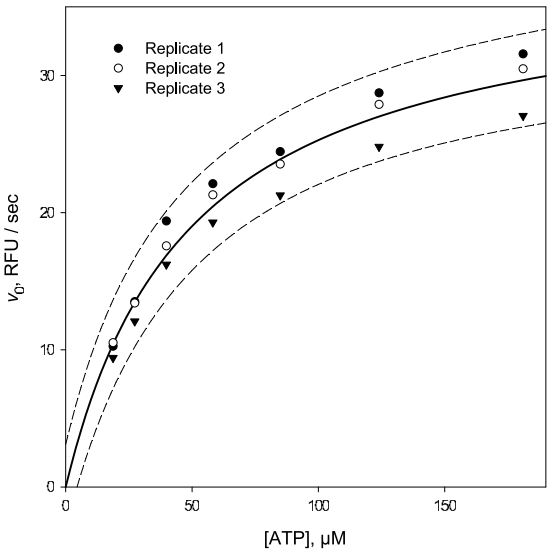


Fig. S 4: Substrate kinetics of ATP: Symbols - initial reaction rates computed from the best-fit values of exponential amplitude A and exponential rate constants k, using Eqn (7). *Solid curve* - least square model curve corresponding to the Michaelis-Menten Eqn (8). *Dashed curves* - inference (prediction) bands at the 95% confidence level. For best fit values of K_{m} and $V_{\max}$ see text.

In the interpretation of the inhibition kinetic results, we utilized this particular K_{m} value to convert all apparent inhibition constants (9) to "true" inhibition constants while assuming that all EGFR inhibitors investigated in this study were strictly *ATP competitive*. Under those assumptions the "true" inhibition constant, K_{i}, can be obtained from the apparent inhibition constant, K_{i}^{*}, as shown in Eqn (10).

$$K_{\mathrm{i}} = \frac{K_{\mathrm{i}}^{*}}{1 + [ATP]_0/K_{\mathrm{m,ATP}}} \tag{10}$$

In all inhibition assays we utilized [ATP] = 800 μM. Thus, the presumed "true" inhibition constants reported in this study are always (1 + 800/50) = 17 times lower than the apparent inhibition constants determined in data fitting.

2. Initial rate kinetics of EFGR inhibitors

In this section we describe the determination of apparent inhibition constants from initial reaction rates. The main purpose was to compare the results with the corresponding values of inhibition constants obtained by the global fit of complete reaction progress curves, and in so doing verify the inhibition constants obtained by either of the two independent method. The initial rates were determined by the least-squares fit on the "early" portion of each inhibition progress curve to an empirical model, represented by the first-order exponential. The reaction rates so obtained were subsequently fit to the Morrison Eqn [8] for tight-binding inhibition.

2.1. Exponential fit of "early" reaction progress

All EGFR inhibitors listed in Table 2 of the main manuscript contain the α,β-unsaturated carboxamide moiety and therefore act as irreversible inhibitors of EGFR. Furthermore, the inhibitory effect for most inhibitors in this study is seen very prominently already at very low inhibitor concentrations comparable with the concentration of the enzyme ("tight binding") [8]. For these reason the complete reaction progress curves ($t_{\max}$ = 30 min) have a complex shape that can only be described by a numerical model represented by a system of differential equations [9]. However, we observed that the initial portion ($t_{\max}$ = 7 min) of every progress curve at any given inhibitor concentration can be successfully fit to the single exponential equation (6). Fig. S5 shows the results or the exponential fit for neratinib. The initial reaction rates form EGFR assay at various inhibition concentrations were computed as the product of exponential amplitudes A and the corresponding first-order rate constants k, as shown in Eqn (7).

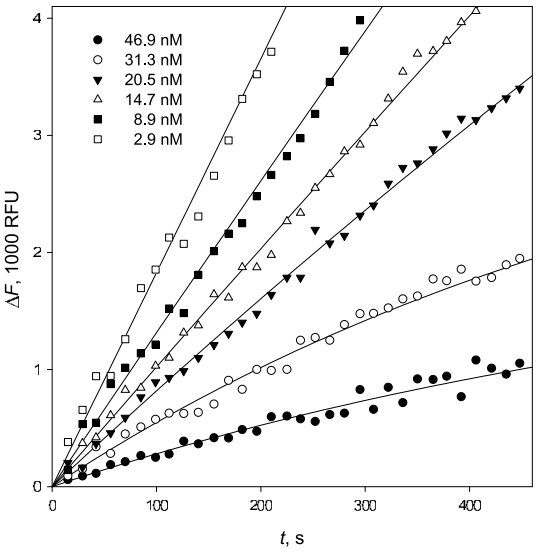


Fig. S 5: Initial rate kinetics of neratinib (one of three replicates): The initial portion of each complete reaction progress curve ($t_{\max}$ = 7 min) was fit to the single-exponential equation (6). Initial reaction rates (see Fig. S1.6) were computed from the exponential amplitudes A and from the first-order rate constants k by using Eqn (7). *Legend:* neratinib concentrations used in each kinetic experiment.

It is crucially important to note that the initial slope of the progress curves shown in Fig. S5 (more precisely, slope of the tangent to each curve at the initial time t = 0) clearly does depend on the inhibitor concentration. For example the initial slope of the best-fit exponential curve drawn through the solid circles ([I] = 46.9 nM) is clearly lower than then initial slope of the best-fit exponential curve drawn through open squares ([I] = 2.9 nM). This fact has a profound mechanistic implication. In particular, the changes in the initial slope clearly indicate that, immediately after the enzyme and inhibitor are mixed, a significant fraction of the enzyme instantaneously forms a significant amount of the non-covalent enzyme/inhibitor complex. In other words, the irreversible inhibition of EFGR by neratinib must proceed in two clearly distinguished mechanistic steps. In

the first reversible step (effectively instantaneous on the timescale of these kinetic experiments) we observe the formation of the non-covalent complex. In the second, much slower and irreversible step, we subsequently observe the formation of the covalent conjugate.

2.2. Determination of apparent inhibition constants

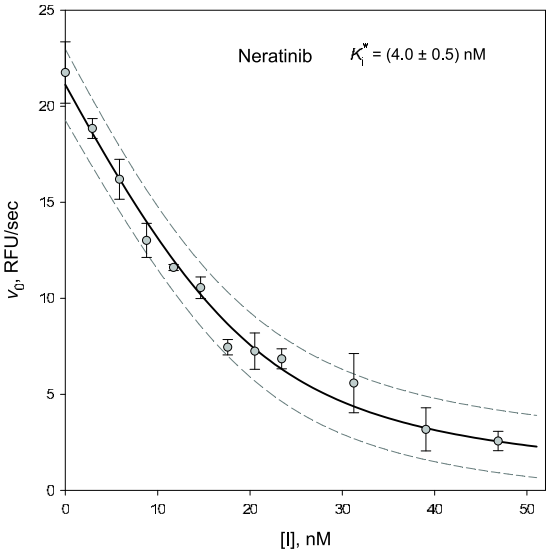


Fig. S 6: Initial rate kinetics of neratinib: *Symbols* - initial reaction rates computed from the best-fit values of exponential amplitude A and exponential rate constants k, using Eqn (7). Each data point is an average from three independent determinations. *Error bars* - standard deviations from replicated (n = 3) initial rate measurements. *Solid curve* - least square model curve corresponding to the Morrison Eqn (11). *Dashed curves* - inference (prediction) bands at the 95% confidence level.

The initial reaction rates, determined by exponential fit of the initial portion of each inhibition progress curve, were fit to the Morrison Eqn (11) to assess the strength of enzyme/inhibitor binding in the initial non-covalent complex. In Eqn (11), V_0 is the uninhibited reaction rate observed in the absence of the inhibitor; $[E]$ is the enzyme concentration; $[I]$ is the inhibitor concentration; and K_i^* is the apparent inhibition constant. The adjustable parameters were V_0 and K_i^*. For all ATP competitive inhibitors, the apparent inhibition constant relates to the "true" inhibition constant as is shown in Eqn (12).

$$v = V_0 \frac{[E] - [I] - K_i^* + \sqrt{\left([E] - [I] - K_i^*\right)^2 + 4[E]K_i^*}}{2\,[E]} \quad (11)$$

$$K_i^* = K_i \left(1 + \frac{[\mathrm{ATP}]}{K_{m,\mathrm{ATP}}}\right) \quad (12)$$

The results of fit are shown in Fig. S6 for neratinib as a representative example. The best-fit value of the apparent inhibition constant was K_i^* = (4.0 ± 0.5) nM. This result corresponds the "true" inhibition constant K_i = (0.24 ± 0.03) nM, according to Eqn (10). It is important to note that the concentration of the peptide substrate was shown to be very much lower than the corresponding Michaelis constant $K_{m,\mathrm{Pep}}$. Thus, the "true" inhibition constants computed by using Eqn (10) above do represent the thermodynamic dissociation equilibrium constant of the instantaneously formed non-covalent complex.

The apparent inhibition constant determined for neratinib by the initial rate method, K_i^* = 4.0 nM, agrees very well with the dissociation equilibrium constant computed from the best-fit values of the association and dissociation rate constant (see Results in the main article), defined as $K_i^* = k_{off}/k_{on}^*$. In three independent kinetic determinations, the k_{off}/k_{on}^* ratio was found to be 3.4 nM, 3.6 nM, and 3.3 nM, respectively. Within the bounds of the formal standard errors, these values agree very well with K_i^* = 4.0 nM determined here from initial rates.

The inhibition constants determined from initial rates, for all irreversible inhibitors of EGFR listed in Table 2 of the main manuscript, are summarized in Table S1.

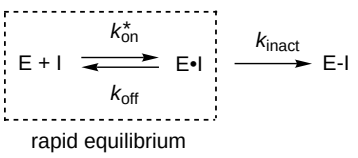
Compound	K_i^*, nM
CI-1033	1.6 ± 0.7
dacomitinib	18 ± 1
afatinib	4.4 ± 0.4
neratinib	4.0 ± 0.5
1	3.2 ± 0.5
CL-387785	180 ± 11
2	63 ± 5
3	84 ± 3
4	2200 ± 100
5	430 ± 30
WZ-4002	340 ± 20

Table S 1: Apparent inhibition constants determined from the fit of *initial reaction rates* to the Morrison Eqn (11) and subsequent conversion of K_i to K_i^* using Eqn (10). See Fig. S6 for a representative example.

3. Global analysis of full reaction progress curves

3.1. Rapid-equilibrium two-step mechanism

In the immediately preceding section, based on the analysis of initial reaction rates, we have established that EFGR is inhibited by α,β-unsaturated compounds listed in Table 2 of the main manuscript in two clearly distinct steps. The first step, essentially instantaneous under our particular experimental conditions, is the reversible formation of the non-covalent enzyme-inhibitor complex (E·I in *Scheme S2*). The second step is the gradual, irreversible formation of the covalent complex (E-I in *Scheme S2*). The dashed-line box shown in *Scheme S2* represents that the formation of the initial non-covalent complex is essentially instantaneous on the time scale of our kinetic experiments.



Scheme S2

Because the formation of the initial complex is essentially instantaneous on the time-scale of our experiments, from our transient kinetic data it is *not possible* for us to determine either the microscopic bimolecular association rate constant for the formation of the reversible enzyme-inhibitor complex, k^*_{on}, or the microscopic dissociation rate constant, k_{off}. Only the ratio $K^*_i = k_{off}/k^*_{on}$ is directly measurable under the given experimental conditions.

3.2. Global fit to an ODE model

DynaFit input script file. The DynaFit [2] script shown in Fig. S7 was used to analyze the time course of EGFR double mutant being inhibited by neratinib. Essentially identical scripts were utilized for all 11 compounds listed in Table 2 of the main manuscript. Each of the 11 compounds was analyzed in three independent experiments (three separate plate-reader plates in the Omnia assay format). All 33 DynaFit scripts and the corresponding data files from 33 independent experiments are available upon request.

The notation "algorithm = differential-evolution" signifies that the global [10] least-square fit of combined progress curves was performed by using the Differential Evolution (DE) algorithm [11]. Briefly, the DE algorithm is an Evolutionary Strategy scheme starting from a population of initial estimates spanning a sufficiently wide range of allowable values. In this case, DynaFit (under its default settings) randomly generated 300 initial parameter estimates spanning 12 orders of magnitude for each of the four rate constants listed in the "[constants]" section of the script. More specifically, the numerical value of each individual rate constant was allowed to vary from 10^{-6} to 10^6. The "[mechanism]" section is a DynaFit representation of the reaction mechanism displayed in *Schemes S1* and *S2*, combined.

All concentrations in the listing shown in Fig. S7 are in micromolar units. The notation "E = 0.02 ? (0.005 .. 0.035)" signifies that the nominal enzyme concentration ([E] = 20 nM) should be treated as one of the adjustable model parameters spanning from 5 to 35 nM. The notation "conc I = 0.03906 ?" (and similarly for other data sets in the global super-set of experimental data shown in Listing S1.1) signifies that all nonzero inhibitor concentrations should also be treated as adjustable model parameter. The notation "Constraints Concentrations = 0.2" signifies that the inhibitor concentrations are allowed to vary by at most 20% of their respective nominal value (in this example [I] = 39.06 nM). The notation "P = 5000 ? (1000 .. 10000)" signifies that the difference molar response coefficient (i.e., the change in fluorescence intensity in Relative Fluorescence Units per micromolar amount of product formed) should be constrained within one order of magnitude. These particular constraints were set on the basis of preliminary experiments. The notation "offset = auto ? (-2000 .. +1000)" signifies that the offset on the fluorescence intensity axis should be optimized, separately for each recorded progress curves, within the indicated bounds.

In accordance with the postulated mechanism, the software internally generated as the least-squares fitting model the system of Ordinary Differential Equations (ODE system) (8) - (13). This ODE system was integrated numerically using the LSODE algorithm [12].

$$\begin{aligned}
d[E]/dt &= -k_{on}[E][I] + k_{off}[E.I] && (13)\\
d[S]/dt &= -k_{sub}[E][S] && (14)\\
d[P]/dt &= +k_{sub}[E][S] && (15)\\
d[I]/dt &= -k_{on}[E][I] + k_{off}[E.I] && (16)\\
d[E.I]/dt &= +k_{on}[E][I] - k_{off}[E.I] - k_{inact}[E.I] && (17)\\
d[E{-}I]/dt &= +k_{inact}[E.I] && (18)
\end{aligned}$$

Results of global fit. The results of fit for three independent kinetic experiments with neratinib are shown in Table S2. The overlay of the best-fit model on the experimental data for one of the three replicates is shown in graphical form in Fig. S8.

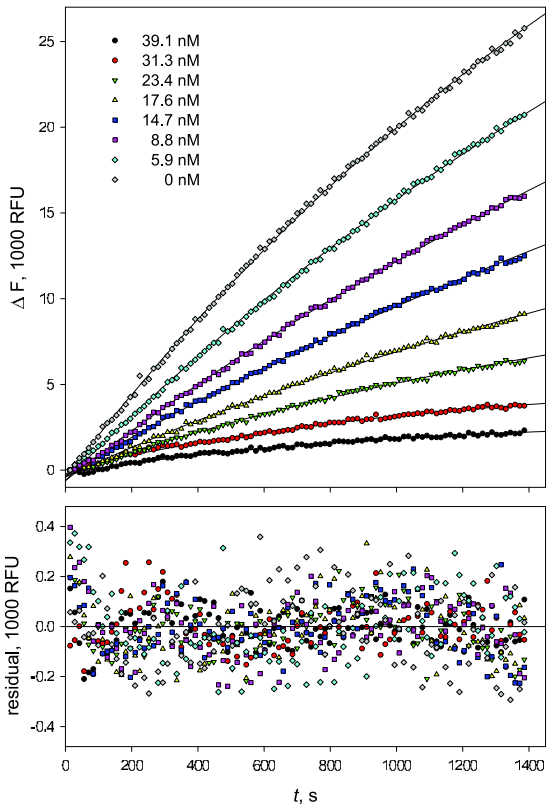


Fig. S 8: Global fit of the reaction progress from inhibition of EGFR-T790M/L858R double mutant by neratinib. The nominal concentrations of neratinib are shown in the inset. The best-fit concentrations are listed in Table S2. *Bottom panel* - residuals of fit.

A fundamentally important question arises regarding the confidence intervals for microscopic rate constants appearing in the combined *Schemes S1* and *S2*. Namely, the question is which of the four microscopic rate constants (if any) can be reliably determined from our particular type experimental data. The answer is unambiguously obtained upon examining the replicated values of k_{sub}, k^*_{on}, k_{off}, and k_{inact} listed in Table S2.

In this context it is very important to note that the "plus or minus" values listed in the right-most column are *not* the formal standard errors arising from nonlinear regression of any particular global data set. Instead the averages and standard deviations for microscopic rate constants listed in Table S2 arose in three entirely independent kinetic experiments, each "experiment" consisting of six to nine reaction progress curves obtained at various inhibitor concentrations pooled together and subjected to global [10] fit.

Thus, the rate constant values in the right-most two columns of Table S2 are averages and standard deviations from the three independent replicates. The controversial [3, 13, 14] formal standard errors of rate constants from nonlinear regression [15, p. 815, Eqn 5.6.4] were entirely ignored in this report.

To assess which particular microscopic rate constant is sufficiently well defined by our experimental data, we reasoned that any particular microscopic rate constant that is well defined will also be well reproduced upon going from one independent set of kinetic experiments to the next replicated set. The conclusions are as follows.

The microscopic rate constant k_{sub} (i.e., the specificity number, $k_{sub} = k_{cat}/K_m$) is very well reproduced across independent replicates, with the coefficient of variation CV approximately equal to 20%. The inactivation rate constant, k_{inact}, is also very well reproduced across all independent replicates (CV $\approx$ 20%, $n = 3$).

In contrast, the bimolecular association rate constant k^*_{on} is not defined by the kinetic data shown in Fig. 8. The dissociation rate constant k_{off} (i.e., the "off rate" constant k_{off}) is not defined by the data, either. However, the rate constant ratio k_{off}/k^*_{on} (i.e., the apparent inhibition constant K^*_i) is very well reproduced across all three replicates (CV $\approx$ 20%, $n = 3$).

In other words, to characterize any given irreversible inhibitor on the basis of our particular type of experimental data, we can only determine k_{inact} and $K_i = k_{off}/k^*_{on}$. We cannot determine the individual values of k_{off} or k^*_{on}. Thus, by implication, neither can we determine at all the value of $K_I = (k_{off} + k_{inact})/k^*_{on}$. As will be shown in a forthcoming report, only the lower limit for both k^*_{on} and k_{off} can be determined under rapid-equilibrium experimental conditions, such as those utilized in our study.

The results shown in Table S2 demonstrate that the association rate constant k^*_{on} is not defined by our type of transient kinetic data. The nominal average value of k^*_{on} is 3.7×10^{10} $M^{-1}s^{-1}$, which would appear to violate the diffusion limit of approximately 10^9 $M^{-1}s^{-1}$. However, the nominal standard error, $\pm 4.6 \times 10^{10}$ $M^{-1}s^{-1}$, is greater than the average value itself. Under such circumstances the "best-fit" value of any rate constant is meaningless, which why we identified k^*_{on} as undefined.

A detailed analysis revealed that any value of k^*_{on} that is greater than 100 $\mu M^{-1}s^{-1}$ will fit all our data sets equally well, for all compounds. Therefore, in the next round of kinetic anal-

Parameter	Low	High	Replicate 1	Replicate 2	Replicate 3	Mean ± Std.Dev.
k_{sub}, $\mu M^{-1}s^{-1}$	10^{-6}	10^6	0.019	0.015	0.019	0.018 ± 0.002
k^*_{on}, $\mu M^{-1}s^{-1}$	10^{-6}	10^6	89000	22000	16000	37000 ± 46000
k_{off}, s^{-1}	10^{-6}	10^6	216.7	45.7	3.2	89 ± 113
K^*_i, nM	–	–	2.4	2.1	2.0	2.2 ± 0.2
k_{inact}, s^{-1}	10^{-6}	10^6	0.0008	0.0011	0.0013	0.0011 ± 0.0002
[E], nM	5	35	17.5	22.4	22.3	20.7 ± 2.8
ϵ_P, RFU/μM	2000	8000	5290	5310	4470	5030 ± 480
[I]#1, nM	31.2	46.9	33.6	32.5	33.1	33.1 ± 0.6
[I]#2, nM	25.0	37.5	25.4	27.0	27.1	26.5 ± 0.9
[I]#3, nM	18.8	28.1	19.2	21.2	21.2	20.5 ± 1.2
[I]#4, nM	14.1	21.1	15.3	17.6	17.9	16.9 ± 1.4
[I]#5, nM	11.7	17.6	12.4	14.7	14.1	13.7 ± 1.2
[I]#6, nM	7.0	10.5	8.4	10.1	10.3	9.6 ± 1.0
[I]#7, nM	4.7	7.0	5.2	5.9	5.5	5.5 ± 0.4
offset#1, RFU	-2000	1000	-186	-21	-198	-135 ± 99
offset#2, RFU	-2000	1000	-95	35	11	-16 ± 69
offset#3, RFU	-2000	1000	-84	1	-296	-126 ± 153
offset#4, RFU	-2000	1000	-298	-390	-405	-364 ± 58
offset#5, RFU	-2000	1000	-145	-177	-361	-228 ± 116
offset#6, RFU	-2000	1000	-460	-387	-606	-485 ± 112
offset#7, RFU	-2000	1000	-454	-632	-608	-565 ± 97
offset#8, RFU	-2000	1000	-558	-617	-407	-528 ± 108

Table S 2: Results of Differential Evolution (DE) global least-squares fit of three independent kinetic experiments with neratinib vs. EGFR-L858R/T790M. The "Low" / "High" columns show the constraints imposed on each given parameters. Each replicate measurement consisted of eight progress curves obtained at the nominal enzyme concentration [E] = 20 nM. The nominal values of seven nonzero inhibitor concentrations are the geometric means of the "Low" and "High" values (see also inset in Fig. S8). The eighth kinetic trace was the control experiment with [I] = 0. The value of K^*_i was computed from the best fit values of microscopic rate constants as k_{off}/k^*_{on}

.

ysis, the rate constant k^*_{on} was held fixed at 100 μM-1s-1 and only the rate constants k_{sub}, k_{off} and k_{inact} were optimized in the regression. For very "tight binding" [8] compounds, but not for others, the nominal enzyme concentration was also optimized in the least squares regression.

Table S3 summarizes the results. As was the case in values listed in Table S2, all standard errors shown in Table S3 are the standard deviations from replicated determinations ($n = 3$), not formal standard errors from nonlinear regression. To compute the K_i values listed in Table 2 of the main manuscript, we have first computed the apparent K^*_i value as the ratio of microscopic rate constants $K^*_i = k_{off}/ k^*_{on}$. Subsequently, assuming that all inhibitors are strictly competitive with ATP, the "true" Ki values were computed using Eqn (10).

The standard deviations from replicates listed in Table S3 are sufficiently small (coefficient of variation lower than 20%) for all three adjustable rate constants, k_{sub}, k_{off}, and k_{inact}.

Strictly speaking, one would expect that the k_{sub} values would be identical across experiments conducted with various inhibitors. The minor variations we do see are probably due to the fact that the enzyme concentration was not exactly identical in each series of experiments ($[E]_0$ and k_{sub} are closely correlated in the regression model).

These results convincingly prove that all three microscopic rate constants treated as adjustable parameters in our differential equations, including k_{sub}, are well determined by our experimental data.

Comparison of K^*_i values from different methods. Fig. S9 shows in graphical form the comparison between the apparent K^*_i values computed by two fully independent methods: (i) from the initial reaction rates obtained by the exponential fit; and (ii) from the global Differential Evolution fit of complete reaction progress curves. The results show that the two methods are in excellent agreement (coefficient of determination $R^2 > 0.99$). The slope of the regression line (0.85) suggests that the apparent K^*_i values computed by progress curve analysis are systematically 15% lower than the apparent K^*_i values obtained by the analysis of initial reaction rates.

These results convincingly prove that, with our newly described data-analytic method, we can reliably determine subnanomolar K_i values even though the enzyme concentration is higher than 10 nM.

4. Biochemical vs. cellular potency correlations

In this section we present a detailed graphical comparison between biochemical potency, as measured by various kinetic parameters (k_{inact}, K_i, and k_{inact}/K_i).

The results displayed in Fig. S10 are drawn from the numerical values listed in Table 2 of the main manuscript. Cellular potency, as measured by inhibition of autophosphorylation in tumor cells, is influenced both by the chemical reactivity of the "warhead" moiety, as measured by k_{inact}, and the dissociation equilibrium constant of the initially formed non-covalent enzyme/inhibitor complex, as measured by K_i* or K_i. This is evidenced by the fact that both k_{inact} and K_i* show significant correlation with the cellular IC_{50}.

```
[task]
data = progress | task = fit | algorithm = differential-evolution

[mechanism]
E + S ---> E + P : ksub
E + I <==> E.I   : kon*     koff
E.I ---> E-I     : kinact

[constants]      | ksub = 1 ?, kon* = 1 ?, koff = 1 ?, kinact = 1 ?
[concentrations] | E = 0.02 ? (0.005 .. 0.035), S = 13
[responses]      | P = 5000 ? (1000 .. 10000)

[data]
directory ./project/kinase/EGFR/inhib/neratinib/Replicate-1/data
sheet sheet.txt
column  3 | offset = auto ? (-2000 .. +1000) | conc I = 0.03906 ?
column  4 | offset = auto ? (-2000 .. +1000) | conc I = 0.03125 ?
column  5 | offset = auto ? (-2000 .. +1000) | conc I = 0.02344 ?
column  7 | offset = auto ? (-2000 .. +1000) | conc I = 0.01758 ?
column  8 | offset = auto ? (-2000 .. +1000) | conc I = 0.01465 ?
column 10 | offset = auto ? (-2000 .. +1000) | conc I = 0.00879 ?
column 11 | offset = auto ? (-2000 .. +1000) | conc I = 0.00586 ?
column 13 | offset = auto ? (-2000 .. +1000) | conc I = 0

[settings] | {Constraints} | Concentrations = 0.2
[end]
```

Fig. S 7: DynaFit [2] script used to analyze the time course of EGFR double mutant being inhibited by neratinib.

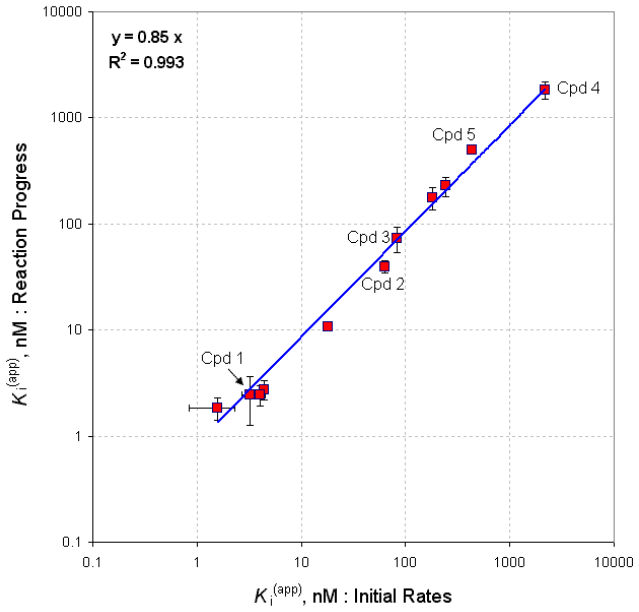


Fig. S 9: Comparison of apparent inhibition constants determined by two different methods. For further details see text.

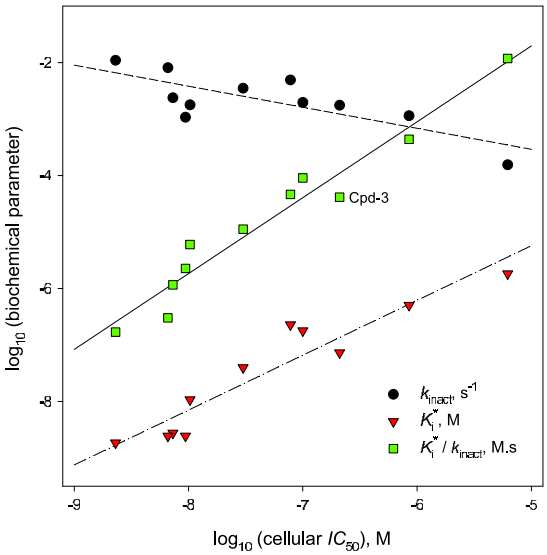


Fig. S 10: Correlation of covalent inhibitor kinetic constants toward EGFR-L858R/T790M with cellular potency (inhibition of EGFR-L858R/T790M autophosphorylation in H1975 tumor cells) using data from Table 2 of the main manuscript.

However, the influence of the initial, non-covalent binding interactions is far stronger than the chemical reactivity of the "warhead". This is made evident by the fact that the coefficient of determination for $\log(k_{\text{inact}})$ vs. $\log(\text{IC}_{50})$ is merely $R^2 = 0.60$, suggesting only weak to moderate correlation. In contrast, the correlation coefficient for $\log(K_{\text{i}})$ is 0.89, suggesting strong correlation.

Interestingly, the strongest correlation ($R^2 = 0.95$) between the biochemical and cellular potency is seen for the ratio $K_{\text{i}}/k_{\text{inact}}$ vs. $\log(\text{IC}_{50})$. In fact if the eleven inhibitors in Table 1 were ordered in terms decreasing $K_{\text{i}}/k_{\text{inact}}$, the resulting ordering would correctly predict the cellular potency (i.e., decreasing cellular IC_{50}) for ten out of eleven inhibitors listed in Table 1. Only compound **3** would appear moderately out of order.

In this context it should be noted that the inverse ratio, $k_{\text{inact}}/K_{\text{i}}$, can be viewed as the lower limit estimate on the microscopic bimolecular association rate constant k_{on}. The microscopic rate constant k_{on} characterizes the "coming together" of the enzyme and the inhibitor, to form the initial non-covalent complex.

In addition to the 11 compounds listed in Table 2 of the main manuscript, we analyzed the correlation between biochemical potency (K_{i} against EGFR-L858R/T790M double mutant) and cellular potency (autophosphorylation in H1975 tumor cells) for 154 cell-permeable, covalent inhibitors spanning six inhibitor scaffolds, see Fig. S11. The coefficient of determination corresponding to the plot in Fig. S11 is $R^2 = 0.72$, which again suggests that initial non-covalent binding contributes very importantly to overall cellular potency.

5. Kinetics of different oxidation states

Table S4 lists the apparent k_{cat} and K_{m} values of different oxidation states of EGFR mutants. *All kinetic parameters listed*

compound	k_{sub}, $\mu\text{M}^{-1}\text{s}^{-1}$	k_{off}, s^{-1}	k_{inact}, s^{-1}	K_{i}^{*}, nM
CI-1033	0.028 ± 0.005	0.19 ± 0.04	0.011 ± 0.0002	1.9 ± 0.4
dacomitinib	0.023 ± 0.002	1.1 ± 0.1	0.0018 ± 0.0001	10.7 ± 0.9
afatinib	0.017 ± 0.004	0.3 ± 0.1	0.0024 ± 0.0003	2.8 ± 0.6
neratinib	0.016 ± 0.002	0.2 ± 0.1	0.0011 ± 0.0002	2.4 ± 0.5
1	0.023 ± 0.004	0.2 ± 0.1	0.008 ± 0.004	2.4 ± 1.2
CL-387785	0.017 ± 0.001	18 ± 4	0.0020 ± 0.0003	180 ± 40
2	0.014 ± 0.004	4 ± 0.5	0.0035 ± 0.0006	40 ± 5
3	0.026 ± 0.003	7 ± 2	0.0018 ± 0.0001	70 ± 20
4	0.021 ± 0.001	180 ± 30	0.00015 ± 0.00002	1800 ± 300
5	0.020 ± 0.004	50 ± 5	0.0011 ± 0.0001	500 ± 40
WZ-4002	0.024 ± 0.001	23 ± 5	0.0049 ± 0.0015	230 ± 50

Table S 3: Results of global fit reaction progress curves (EGFR-L858R/T790M) to the system of differential equations (13)–(18). *The plus-or-minus values are standard deviations from averaging three replicated, entirely independent experiments.* For further details see text.

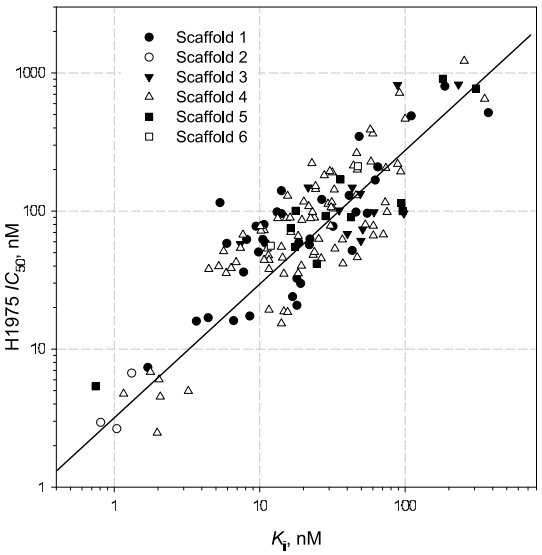


Fig. S 11: Correlation of dissociation constants of the initial non-covalent enzyme/inhibitor complex, as measured by the K_i values for EGFR-L858R/T790M double mutant, with cellular potency (inhibition of EGFR-L858R/T790M autophosphorylation in H1975 tumor cells) for 154 compounds spanning six structural classes.

as with respect to ATP, not with respect to the peptide substrate. The peptide substrate in these determinations was always present at concentrations very much lower than the corresponding Michaelis constant. In fact, the apparent k_{cat} and K_m values with respect to the peptide substrate could not be determined for practical reasons (see section *Substrate kinetics*).

6. Inhibitor Pharmacology: Oxidized EGFR-Cys$_{797}$

Table S5 shows the potency values for inhibitors of oxidized EGFR-L858R single mutant in biochemical and cellular assays. Cellular potency was quantified by determining the inhibition of EGFR autophosphorylation in H3255 tumor cells. Table S6 shows similar results for the EGFR-L858R/T790M double mutant. In this case cellular potency is quantified as inhibition of EGFR autophosphorylation in H1975 tumor cells. The results show that EGFR-Cys$_{797}$ oxidation (-SH, unoxidized, -SO_2H sulfinylated, -SSG, glutathiolated) can profoundly affect inhibitor affinity depending on the inhibitor characteristics and EGFR context.

7. Surface plasmon resonance (SPR) binding studies

To evaluate the energetic contribution of the Cys$_{797}$-Michael Acceptor interaction to overall potency, reversible EGFR inhibitor affinities to EGFR were determined by surface plasmon resonance and compared to the inhibitor affinities (K_i) of the corresponding covalent inhibitors determined by either irreversible or reversible inhibition enzyme kinetics.

SPR studies were carried out on a Biacore 3000 instrument at 25°C in 150mM NaCl, 25mM HEPES, pH 7.2, 5mM $MgCl_2$, 5% glycerol, 1% DMSO, 0.5 mM TCEP, 0.005% P20. EGFR proteins were immobilized on a CM5 sensor chip by standard amine coupling, in 10 mM sodium acetate pH 5.0 (1 M EGFR). Injections were made using the Kinject mode at 50 μL/min with a 200-1200s compound dissociation time. Compound injections were referenced to a blank surface and by a buffer blank. Data analysis and fitting to a simple 1:1 kinetic model was performed using the Scrubber2 software (BioLogic Software, Pty., Australia).

The results are summarized in Table S7 for wild-type enzyme and in Table S8 for the L858R/T790M double mutant.

8. Mass spectrometric characterization of EGFR

8.1. Mass spectrometry methods

Intact mass analysis of modified EGFR protein was measured by electrospray ionization on an Agilent (San Jose, CA) 6210 time-of-flight mass spectrometer (TOF-MS) coupled to an Agilent 1200 LC. EGFR protein samples, 500 ng/10 μL, were injected onto an Acquity UPLC BEH300 C4 2.1 $\times$ 100 mm column (Waters, Bedford, MA.). At a flow rate of 0.5 mL/min, a six-minute mobile phase gradient was constructed using 0.1% formic acid (Solvent A) and acetronitrile (Solvent B). In the first two minutes of the gradient, acetonitrile (Solvent B) was ramped up to 40%, during which the column elution was sent to waste. Subsequently, the LC stream was directed into the MS, and solvent B was increased to 60% for the next three minutes for elution of EGFR protein. In the final minute of the gradient, the column was exposed to 90% acetonitrile and then allowed re-equilibrated to 2% Solvent A. The protein was eluted off the column into the MS-TOF, which was set to detect a mass range from 600-2000 m/z. The MS fragmentor was set at 200 V, and the skimmer at 140 V. Once eluted, the protein spectra were extracted from the eluted chromatography peaks in Agilent Qualitative Analysis software, and subsequently deconvoluted with the Maximum Entropy algorithm. The deconvolution was run over a mass range of 10-100 kDa, with a 1 Dalton mass step, and 20 signal/noise threshold.

For identification of modified residues, "bottom-up" mass spectrometry was employed, using nanoLC-MS. For "in-gel" digestion, EGFR protein was first denatured in 8M urea (30 min, room temp) and then alkylated with 10 mM iodoacetamide (2 hour, room temp, in the dark). Subsequently, 10 μL of sample (0.5 μg/μL) was loaded into wells on a non-reducing SDS-PAGE gel, which was Coomassie blue stained. EGFR gel bands were excised, cut and destained twice in 60/40 acetonitrile/water for 1 hour at room temp. For the subsequent proteolysis, the gel slices were first washed in 5% formic acid, before addition of 50 ng of pepsin (Protea, Morgantown, WV.) and incubated for 16 hours at 37°C. The resultant EGFR digest was analyzed on a LTQ mass spectrometer (Thermo Fisher Scientific) equipped with a Michrom Captive Spray ionization source, and coupled to a Proxeon nanoLC. Sample was injected at 5uL onto a Reprosil ProteCol Trap C18-AQ (SGE Analytical Sciences), for initial desalting with 20 μL of 0.1% formic acid solution. Analytical separation was performed using a Halo ES-C18 column (Michrom BioResources, Auburn, CA.), 0.2x150 mm, running at a flow rate of 1.8 μL/min. The LC gradient consisted of a 35-minute gradient from 2-40% acetronitrile with

0.1% (v/v) formic acid to elute peptides into the mass spectrometer. Mass spectrometry data was collected in "triple-play" mode, consisting of a full scan at 400-2000 m/z for selection of the three most predominant eluting peptides, followed by a high resolution ultra-zoom scan, and MS-MS of isolated peaks of interest. MS-MS spectra were collected in an isolation window of 3 m/z and a collision activation energy (CAD) of 35 eV. Data was then processed in Agilent Spectrum Mill rev. 4.0, for peptide identification. The EGFR mutation protein sequences were appended to the Spectrum Mill in-house database, and peptides were search directly against the prescribed EGFR sequence. Also, peptides were searched with variable methionine oxidation, cysteine oxidation and cysteine glutationylation. Precursor mass tolerance was set at 1.0 Da, and 0.7 Da for the fragment ion. Peptides were validated by fixed thresholds of forward-reverse scores < 8, with spectrum intensities (SPI) < 70. The search parameters were set for up to three missed cleavage sites.

8.2. Mass spectrometry results

Localization of EGFR oxidation was performed by LC-MS-MS analysis of the EGFR pepsin digest. Pepsin was the preferred protease because it provided greater than 89% sequence coverage for both the single activating mutation (L858R) and double mutant EGFR (T790M, L858R), and most importantly allowed for observation all six cysteine residues. LC-MS-MS analysis of the oxidized single and double mutant EGFR protein identified Cys_{797} as the site of oxidation. Shown in Fig. S12 is both the intact mass MS spectrum of the oxidized double mutant EGFR (T790M, L858R), and the subsequent MS-MS spectra for the pepsin generated peptides. Verification of Cys_{797} oxidation to its sulfinic acid form was verified from MS-MS fragmentation of the proteolytic peptide from residues 782-805, from which Spectrum Mill data analysis confidently assigned modification of Cys_{797} (Score = 18.40, Fwd-Rev Score = 18.40). It should be noted that both the cysteine -sulfinic (SO_2) and -sulfonic (SO_3) acid forms were observed in MS-MS analysis, which is not surprising given there was indication of a minimal portion (<10%) of SO_3 oxidation in the intact mass analysis. EGFR-Cys_{797} was identified as the only EGFR residue with an altered oxidation state for both the EGFR-L858R and EGFR-L858R/T790M H_2O_2 oxidized samples (Fig. S12). As with H_2O_2 oxidation, glutathiolation of EGFR-Cys_{797} was the only modification site for both observed for both EGFR-L858R and EGFR-L858R/T790M. Shown in Fig. S13 is the MS-MS spectrum for the glutathiolated peptide spanning residues 788-800 of double mutant of EGFR, from which Spectrum Mill data analysis confidently assigned modification of Cys_{797} (Score = 11.70, Fwd-Rev Score = 11.70).

9. Molecular modeling

The active site orientation of the inhibitor Michael Acceptor to the reactive cysteine is defined by the inhibitor scaffold. Molecular modeling (Figure S14) predicts the binding mode of compound **1** (magenta) and WZ4002 (yellow) is very different for the EGFR L858R mutant. The distance is calculated between C(ligand)-S(Cys797) (yellow residue).

We postulate that sulfinylation of EGFR-Cys_{797} may enhance inhibitor affinity by creating a positive interaction between the

	L858R			L858R/T790M			w.t.
Cys_{797} →	-SSG	-SO_2H	-SH	-SSG	-SO_2H	-SH	-SH
k_{cat}, s^{-1}	6.8±0.2	7.6±0.2	6.9±0.4	14.7±0.6	17.8±2.8	15.4±1.3	3.5±0.3
K_m, μM	1120±30	280±20	260±10	340±30	130±10	110±1	110±3
k_{cat}/K_m, $\mu M^{-1}s^{-1}$	0.0061±0.0002	0.027±0.002	0.027±0.002	0.044±0.005	0.14±0.02	0.14±0.01	0.078±0.009

Table S 4: Kinetic analysis of EGFR and oxidized-EGFR (-SH, unoxidized, -SO_2H sulfinylated, -SSG, glutathiolated) catalysis for L858R and L858R/T780M variants using the Ac-EEEEYIIV-NH2 peptide. The S-glutathiolation did not reduce the fraction of functional active sites for either protein but sulfinylation did by 50%. EGFR protein was highly pure but were normalized for the fraction of functional active sites to provide the most accurate kinetic constants.

	biochemical K_i, nM			cellular
compound	-SSG	-SO_2H	-SH	IC_{50}, nM
CI-1033	4.2 ± 0.6	4.7 ± 0.3	0.4 ± 0.1	0.72 ± 0.36
dacomitinib	21 ± 2	1.0 ± 0.1	0.7 ± 0.1	0.83 ± 0.04
1	6.7 ± 0.9	7.0 ± 1.1	0.4 ± 0.1	1.4 ± 0.2
afatinib	13 ± 1	1.7 ± 0.2	0.6 ± 0.1	1.2 ± 0.1
6	29 ± 3	1.4 ± 0.1	1.1 ± 0.1	3.3 ± 0.0
erlotinib	7.4 ± 0.8	5.0 ± 1	2.7 ± 0.1	18 ± 6
gefitinib	46 ± 4	1.2 ± 0.1	1.5 ± 0.2	5.2 ± 2.4
lapatinib	3.5 ± 0.2	0.7 ± 0.1	0.3 ± 0.1	20 ± 3
WZ4002	5100 ± 300	14900 ± 500	13 ± 1	86 ± 21
WZ4003	7700 ± 300	3700 ± 40	1100 ± 100	>10000
8	ND	ND	0.83 ± 0.09	ND
9	ND	ND	3.1 ± 0.2	ND

Table S 5: Potency values for inhibitors of oxidized EGFR-L858R single mutant in biochemical and cellular assays. ND = not determined.

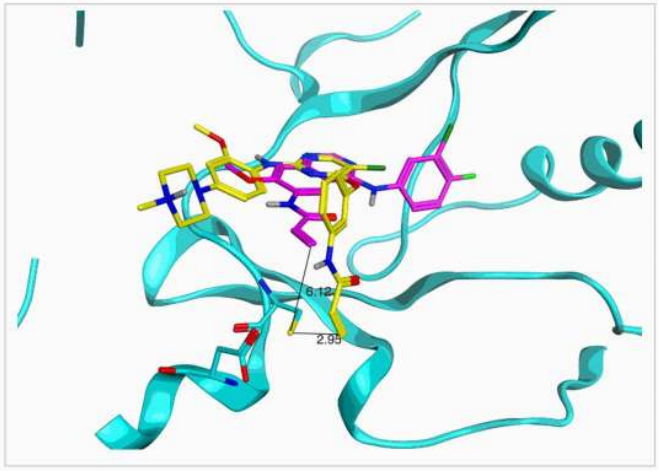


Fig. S 14: Molecular modeling comparison of compound **1** (magenta) and WZ4002 (yellow). For details see text.

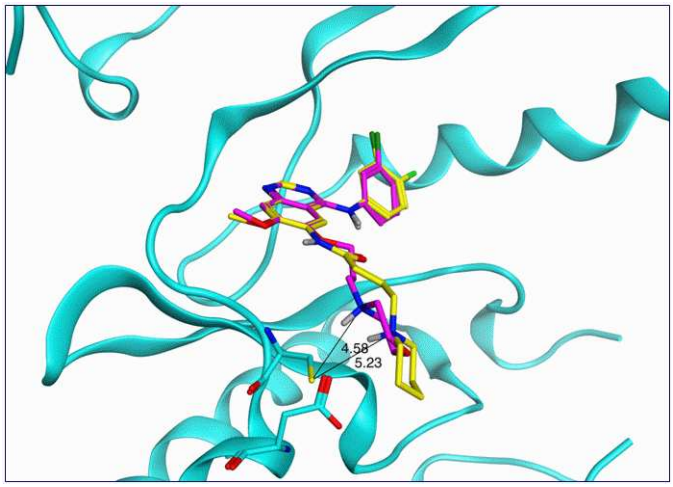


Fig. S 15: Molecular modeling of gefitinib (magenta, PDB ID 2ITZ) and dacomitinib (yellow, docked) in the unoxidized form of EGFR-L858R. For details see text.

oxidized cysteine residue S=O group and an inhibitor basic substituent. Molecular modeling studies (Fig. S15) predict a binding mode of gefitinib (magenta, PDB ID 2ITZ) and dacomitinib (yellow, docked) in the unoxidized form of EGFR-L858R with a short N(ligand)-S(Cys_{797}) distance (4.58 - 5.23 Å). With sulfinylated EGFR-Cys_{797}, the cysteine S=O and the inhibitor NH group are expect to be at a good interaction distance.

Histograms of hydrogen bond length distributions of NH(ligand) and S=O(Cys_{797}) in Cambridge Structure Database (Fig. S16) suggest the mean N–O distance of 2.94 Å (*Top Panel*) and the mean N–S distance of 4.07 Å (*Bottom Panel*).

10. Detailed experimental methods

10.1. Expression and purification of EGFR proteins

The cDNA encoding the core domain (residues 696-1022) of human EGFR double mutant (L858R/T790M) was synthesized and cloned into pKRIC-N3 using the BamHl and HindIII sites. Similarly, the cDNA encoding the single mutant (L858R) core domain (residues 695-1022) was cloned into pKRIC-N5 and the cDNA encoding the WT core domain with the juxtamembrane region (residue 668-1022) was cloned into pKRIC-N3.

Inhibitor	biochemical K_i, nM			cellular IC_{50}, nM
	-SSG	-SO_2H	-SH	
CI1033	16 ± 2	21 ± 2	0.2 ± 0.1	2.3 ± 0.5
dacomitinib	49 ± 6	2.4 ± 0.6	0.9 ± 0.1	10 ± 1
1	36 ± 1	55 ± 8	0.3 ± 0.1	6.6 ± 0.2
afatinib	20 ± 2	2.1 ± 0.2	0.2 ± 0.1	7.3 ± 1.1
6	82 ± 2	2.2 ± 0.4	29 ± 2	100 ± 0
erlotinib	80 ± 7	47 ± 2	54 ± 1	5200 ± 300
gefitinib	420 ± 50	5.5 ± 0.3	22 ± 3	4300 ± 2300
lapatinib	210 ± 10	210 ± 20	120 ± 3	7400 ± 100
WZ4002	2200 ± 200	1400 ± 200	16 ± 4	78 ± 25
WZ4003	1700 ± 100	870 ± 10	340 ± 10	>10,000
8	ND	ND	ND	≈660
7	ND	ND	ND	1800 ± 300

Table S 6: Potency values for inhibitors of oxidized EGFR-L858R/T790M double mutant in biochemical and cellular assays. ND = not determined.

unreactive analog	k_{on}, $\mu M^{-1}s^{-1}$	k_{off}, s^{-1}	K_d, nM	K_i, nM	reactive analog	K_i, nM
7	0.9	0.00046	0.51	0.8	CI-1033	0.093
8	0.84	0.0015	1.8	ND	dacomitinib	0.16
9	1.6	0.0022	1.4	ND	1	0.18
WZ4003	0.021	0.013	620	ND	WZ4002	28

Table S 7: Biacore results vs. biochemical inhibition potency: wild-type EGFR. ND = not determined. For further explanation see text.

Both the double mutant (DM, L858R/T790M) and WT constructs contain an N-terminal his tagged GST followed by a TEV (Tobacco Etch Virus) cleavage site and the single mutant (L858R) construct contains an N-terminal his tag followed by a TEV cleavage site. pKRIC-N3 and pKRIC-N5 are modified insect cell expression cloning vector pFastbac I. Therefore, they are only used in the Bac-to-Bac method for recombinant baculovirus production. Sf9 cells were used to generate recombinant baculovirus and to scale up expression for WT construct and sf21 cells were used to scale up expression for L858R/T790M and L858R constructs.

Specifically, sf9 or sf21 cells were infected with baculovirus based on cell density. Forty eight hours (for L858R and WT) or seventy two (for L858R/T790M) hours post infection, cells were harvested by centrifugation and the cell pellet was frozen and stored at -80 C. To purify the protein, a cell pellet was resuspended in a lysis buffer containing 100 mM HEPES, pH 8.0; 150 mM NaCl; 5mM $MgCl_2$; 10% glycerol; 20 mM imidazole; 4mM TCEP; and protease inhibitor tablets (Roche). Cells were lysed by stirring the suspension at 4 C for 45min and the cell lysate was centrifuged at 10,000g-20,000g for 60min at 4 C. The supernatant was collected by passing through four layers of cheesecloth. The tagged protein was then purified by batch binding with Invitrogen Probond resin for two hours at 4 C with slow rotation using a Roto-Shaker Genie, followed by washing the bound resin with the lysis buffer and then eluting with the lysis buffer containing 300mM imidazole. The untagged and non-phosphorylated protein was obtained by overnight dialysis at 4 C in the presence of TEV protease and lambda phosphatase in a dialysis buffer containing 25 mM HEPES, pH 8.0; 150 mM NaCl; 5 mM $MgCl_2$; 10% Glycerol; 2 mM TCEP; and 1 mM MnCl2, followed by nickel reverse chromatography. The untagged protein was further purified by Superdex75 size exclusion chromatography using an SEC buffer containing 25 mM HEPES, pH 8.0; 50-100mM NaCl; 5 mM $MgCl_2$; 10% glycerol; 4 mM TCEP. The peak fractions were pooled based on the activity and purity (SDS-PAGE analysis). Pools were concentrated to a desired concentration (2-3 mg/mL). The concentra-

unreactive analog	k_{on}, μM^{-1}s^{-1}	k_{off}, s^{-1}	K_d, nM	K_i, nM	reactive analog	K_i, nM
7	0.60	0.0300	50	37	CI-1033	0.22
8	0.45	0.0088	20	ND	dacomitinib	0.86
9	0.22	0.0240	110	ND	1	0.26
WZ4003	0.021	0.0082	390	340	WZ4002	16

Table S 8: Biacore results vs. biochemical inhibition potency: EGFR-L858R/T790M double mutant. ND = not determined. For further explanation see text.

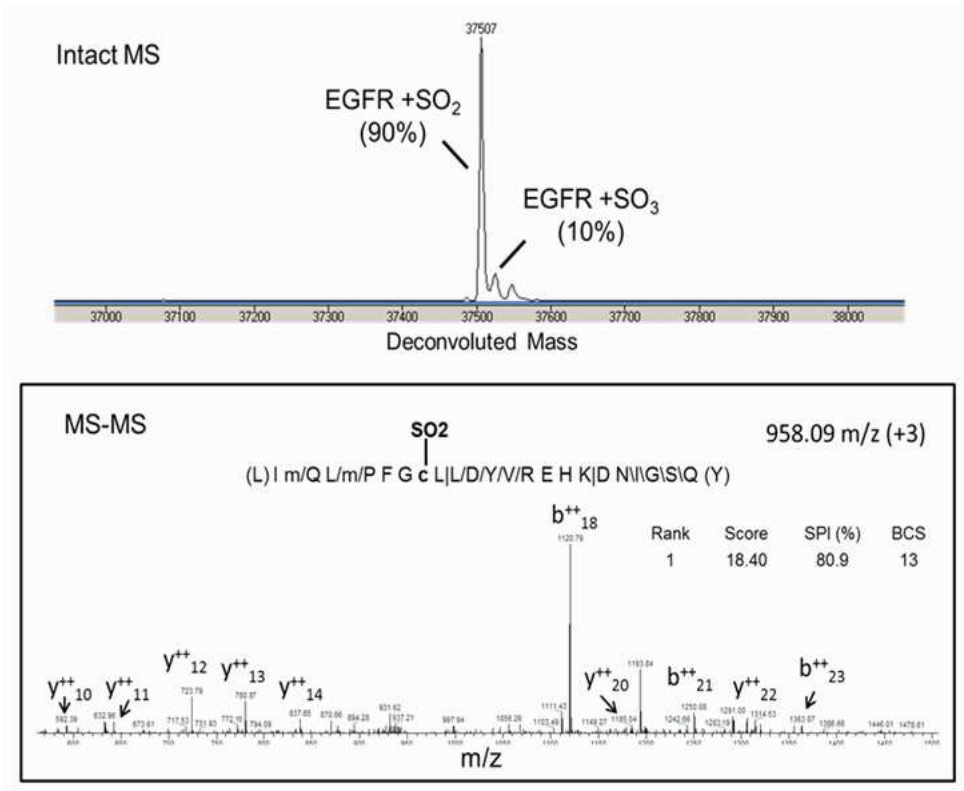


Fig. S 12: Intact mass measure of the oxidized double mutant EGFR (L858R, T790M), indicating 90%, by peak area, in the SO_2-oxized form. Subsequent pepsin digest of this oxidized protein, generated a peptide (residues 782-805) from which MS-MS identification of SO_2 on cysteine-797 is evident. Similar results were obtained for EGFR (L858R).

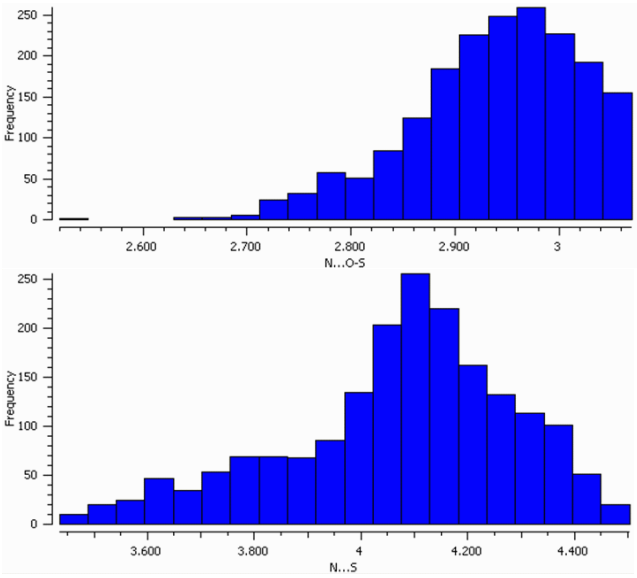


Fig. S 16: Histograms of hydrogen bond length distributions in Cambridge Structure Database. *Top Panel*: N–O distance; *Bottom Panel*: N–S distance. For details see text.

tion was measured using both Abs280 and colorimetric assay using BSA as standard (Biorad coomassie plus protein assay reagent). The protein purity was checked by SDS-PAGE with all proteins having >95% purity. The identity of protein was confirmed by intact mass spectroscopy. For WT protein purification, the lysis and nickel elution buffers had 200 mM NaCl and the dialysis and size exclusion chromatography buffers had 300 mM NaCl.

10.2. Determination of active enzyme concentration

Enzyme concentration for the non-oxidized EGFR mutants were determined by conducting the fluorometric covalent inhibitor assay in the presence of varying amounts of dacomitinib at concentrations ranging from one-tenth to two-fold of the nominal enzyme concentration. Assay conditions were similar to those used to test covalent inhibitors (see section *Fluorometric enzyme assays*), however EGFR (50 nM) and inhibitor were incubated for 20 minutes in the absence of ATP. The assay was initiated by the addition of 1.1 mM ATP.

The initial velocity was plotted against the ratio of [inhibitor]/[enzyme]. A straight line was fit through the data points which represented reactions that were not fully inhibited, and the intercept on the horizontal axis yielded the fraction of enzyme with active sites capable of turnover. L858R/T790M had 96% of the added protein capable of inhibition by this active conformation-binding inhibitor dacomitinib. L858R and WT had a smaller fraction of the active sites available to this inhibitor (53% and 23% respectively).

L858R and WT EGFR proteins were inhibited by the nonactive conformation binding inhibitor lapatinib at varied concentrations. The initial rates were fit to the Morrison equation (11), where the active enzyme concentration $[E]_0$ was treated as one of the optimized parameters along with the apparent inhibition constant K_i^*. The best-fit value of $[E]_0$ was greater than 95% of the nominal enzyme concentration.

The same method was used to estimate the concentration of active sites for EGFR mutants oxidized at Cys_{797} (which are incapable of reactivity with covalent inhibitors). S-Glutathiolated EGFR-L858R/T790M has 100% of the active sites functional, while the sulfinylated form was 50% of the active sites functional. EGFR-L858R protein has 50% of the active sites functional after S-glutathiolation and 25% of active sites available after sulfinylation.

10.3. UV/Vis spectrophotometric enzyme assays

A spectrophotometric coupled enzymatic assay was used to monitor the kinase-catalyzed production of ADP from ATP that accompanies phosphoryl transfer to a phosphoacceptor peptide substrate Ac-EEEEYIIV-NH2. Activity was monitored by coupling NADH oxidation to the regeneration of ATP from product ADP through the action of pyruvate kinase (PK) and lactate

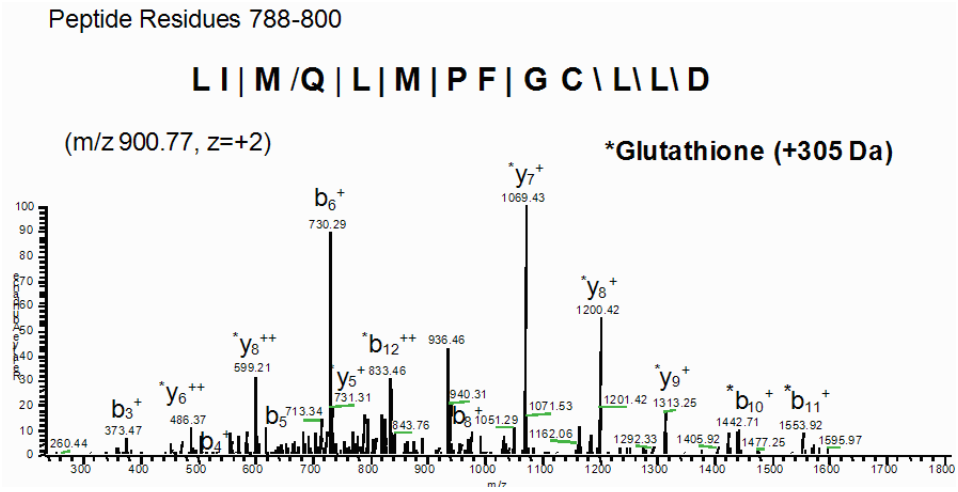


Fig. S 13: MS-MS data from precursor peptide residues 788-800, LIMQLMPFGCLLD, (m/z 900.77, z=+2) confirms glutathiolation on Cys_{797} from glutathiolated EGFR double mutant (L858R, T790M). Similar results were obtained for EGFR (L858R).

dehydrogenase (LDH). NADH conversion to NAD+ was monitored by the decrease in absorbance at 340 and 25^{deg}C.

Typical reaction solutions contained 2 mM phosphoenolpyruvate, 0.27 mM NADH, 12 mM free $MgCl_2$, 1 mM DTT, 2 mM peptide (the solubility limit), 10 units/mL PK, 10 units/mL LDH, 150 mM NaCl, and 0.01% Tween-20 in 50 mM HEPES pH 7.5. ATP concentration was dependent upon the assay performed (Table S8.1). Reactions were performed either at 200 μL volumes in a quartz cuvette on a Beckman DU800 (kinetic parameter determination) or at 100 μL volumes in a clear-bottom, half-area 96-well plate on a Tecan Safire in absorbance mode (inhibitor affinity determination).

Kinetic parameters for the phosphoryl transfer reaction, the turnover number, $k_{cat,ATP}$, and the Michaelis constant for ATP, $K_{m,ATP}$, respectively, were determined using the UV/Vis spectrophotometric coupled assay by varying [ATP] around that which produced half maximal velocity for each EGFR mutant (oxidized and non-oxidized). Reactions were initiated by the addition of 30 to 120 nM EGFR-L858R/T790M or EGFR-L858R (final concentration) for both the oxidized and non-oxidized species.

The initial portion of each reaction progress curve was fit to the straight-line model, to determine the initial rate v, as the slope. To determine the kinetic parameters $k_{cat,ATP}$ and $K_{m,ATP}$, the initial rates so obtained were fit to Eqn (1), where $[E]_0$ is the enzyme concentration independently determined by active-site titration.

Inhibitor affinity (K_i) determinations were performed by fitting initial rate data to the Morrison Eqn (11). Apparent inhibition constants determined by nonlinear regression were converted to "true" inhibition constants by using Eqn (10), assuming strictly ATP-competitive inhibition mode. A liquid handling robot (BiomekFX) was used to make 1:2 inhibitor dilutions from DMSO stocks (2% final DMSO concentration), add inhibitor to the reaction mixture, and initiate the reaction by the addition of ATP substrate. Enzyme and ATP concentrations for each EGFR mutant (oxidized and non-oxidized) are found in Table S9.

Enzyme form	[ATP], mM	[E], nM
EGFR-L858R	1.10	30
Cys_{797} sulfinylated[(a)]	1.10	30
Cys_{797} glutathiolated[(b)]	1.40	28
EGFR-L858R/T790M	0.33	10
Cys_{797} sulfinylated[(a)]	0.33	15
Cys_{797} glutathiolated[(b)]	1.01	10

[(a)] Oxidized to the sulfinic acid at Cys_{797}.
[(b)] Oxidized to the glutathione adduct at Cys_{797}.

Table S 9: Enzyme and ATP concentrations utilized in UV/Vis spectrophotometric assays.

10.4. Fluorometric enzyme assays

Fluorometric assays were also performed by using the Omnia continuous fluorometric kinase assay system (Invitrogen, Carlsbad, California) using peptide Y-12, a fluorogenic tyrosine phosphoacceptor peptide modified with a chelation-enhanced sulfonamide-oxine fluorophore (cSx) coupled to a cysteine residue (Ac-EEEEYI(cSx)IV-NH2). Phosphopeptide formation was monitored in 50 μL reactions in 96-well plates with a Tecan Safire II microplate reader in fluorescence mode using 360 nm excitation and 485 nm emission wavelengths. Reactions were comprised of 12 mM free $MgCl_2$, 1 mM DTT, 13 μM peptide-cSx, 150 mM NaCl, and 0.01% Tween-20 in 50 mM HEPES pH 7.5. ATP concentrations were dependent upon the assay performed (Table S8.1). Reactions were initiated by the addition of 30 nM EGFR-L858R or 20 nM EGFR-L858R/T790M (final concentration).

The apparent $K_{m,ATP}$ under fluorometric assay conditions for each non-oxidized EGFR mutant was determined by varying [ATP] around that which produced half maximal velocity. The Michaelis-Menten equation (14) was fitted to the data. Apparent $K_{m,ATP}$ for EGFR-L858R, EGFR-L858R/T790M, EGFR-WT were determined to be 114, 50, and 34 M respectively.

Detailed kinetic analysis of covalent inhibitors acting on EGFR mutants not oxidized at EGFR-Cys_{797} was also performed by using the fluorometric assay. Determinations were made from a series of progress curves as a function of inhibitor concentration varied around the K_i^* for the reversible binding portion of the reaction. Inhibitor was added from stock dilutions in DMSO (2% final DMSO concentration) to reactions containing 5 mM or 0.8 mM ATP for EGFR-L858R and EGFR-L858R/T790M assays, respectively.

11. Structures of nonreactive inhibitors

Figure 17 displays the chemical structures of nonreactive EGFR inhibitors utilized in this study. Corresponding reactive analogues are shown in parentheses.

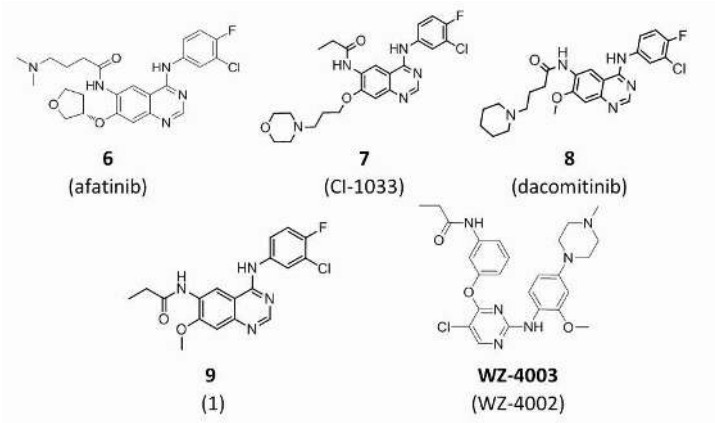


Fig. S 17: Structures of nonreactive compounds used in this study.

References

[1] K. B. Burnham, D. R. Anderson, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd Edition, Springer-Verlag, New York, 2002.

[2] P. Kuzmič, DynaFit – A software package for enzymology, Meth. Enzymol. 467 (2009) 247–280.

[3] K. A. Johnson, Fitting enzyme kinetic data with KinTek Global Kinetic Explorer, Meth. Enzymol. 267 (2009) 601–626.

[4] P. Mendes, H. Messiha, N. Malys, S. Hoops, Enzyme kinetics and computational modeling for systems biology, Meth. Enzymol. 467 (2009) 583–599.

[5] I. H. Segel, Enzyme Kinetics, Wiley, New York, 1975.

[6] D. M. Bates, D. G. Watts, Nonlinear Regression Analysis and its Applications, John Wiley & Sons, New York, 1988.

[7] I. Brooks, D. G. Watts, K. K. Soneson, P. Hensley, Determining confidence intervals for parameters derived from analysis of equilibrium analytical ultracentrifugation data, Methods Enzymol. 240 (1994) 459–478.

[8] J. F. Morrison, Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors, Biochim. Biophys. Acta 185 (1969) 269–286.

[9] S. Szedlacsek, R. G. Duggleby, Kinetics of slow and tight-binding inhibitors, Methods Enzymol. 249 (1995) 144–180.

[10] J. M. Beechem, Global analysis of biochemical and biophysical data, Meth. Enzymol. 210 (1992) 37–54.

[11] K. V. Price, R. M. Storm, J. A. Lampinen, Differential Evolution - A Practical Approach to Global Optimization, Springer Verlag, Berlin - Heidelberg, 2005.

[12] A. C. Hindmarsh, ODEPACK: a systematized collection of ODE solvers, in: R. S. Stepleman, et al. (Eds.), Scientific Computing, North Holland, Amsterdam, 1983, pp. 55–64.

[13] K. A. Johnson, Z. B. Simpson, T. Blom, Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data, Anal. Biochem. 387 20–29.

[14] K. A. Johnson, Z. B. Simpson, T. Blom, FitSpace Explorer: An algorithm to evaluate multidimensional parameter space in fitting kinetic data, Anal. Biochem. 387 (2009) 30–41.

[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes: The Art of Scientific Computing, 3rd Edition, Cambridge University Press, Cambridge, 2007.